

Video recording — 12,65 mm type D-11 format —

Part 2: Picture compression and data stream

The European Standard EN 62356-2:2004 has the status of a
British Standard

ICS 33.160.40

National foreword

This British Standard is the official English language version of EN 62356-2:2004. It is identical with IEC 62356-2:2003.

The UK participation in its preparation was entrusted to Technical Committee EPL/100, Audio, video and multi-media systems, which has the responsibility to:

- aid enquirers to understand the text;
- present to the responsible international/European committee any enquiries on the interpretation, or proposals for change, and keep the UK interests informed;
- monitor related international and European developments and promulgate them in the UK.

A list of organizations represented on this committee can be obtained on request to its secretary.

Cross-references

The British Standards which implement international or European publications referred to in this document may be found in the *BSI Catalogue* under the section entitled “International Standards Correspondence Index”, or by using the “Search” facility of the *BSI Electronic Catalogue* or of British Standards Online.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

Compliance with a British Standard does not of itself confer immunity from legal obligations.

Summary of pages

This document comprises a front cover, an inside front cover, the EN title page, pages 2 to 55 and a back cover.

The BSI copyright notice displayed in this document indicates when the document was last issued.

Amendments issued since publication

Amd. No.	Date	Comments

This British Standard was published under the authority of the Standards Policy and Strategy Committee on 1 February 2005

© BSI 1 February 2005

ISBN 0 580 45327 8

**Video recording –
12,65 mm type D-11 format
Part 2: Picture compression and data stream
(IEC 62356-2:2003)**

Enregistrement Vidéo –
Format 12,65 mm de type D11
Partie 2: Flux de données et compression
d'image
(CEI 62356-2:2003)

Videoaufzeichnung –
D-11-Format mit 12,65 mm
Teil 2: Bildkompression und Datenstrom
(IEC 62356-2:2003)

This European Standard was approved by CENELEC on 2004-09-01. CENELEC members are bound to comply with the CEN/CENELEC Internal Regulations which stipulate the conditions for giving this European Standard the status of a national standard without any alteration.

Up-to-date lists and bibliographical references concerning such national standards may be obtained on application to the Central Secretariat or to any CENELEC member.

This European Standard exists in three official versions (English, French, German). A version in any other language made by translation under the responsibility of a CENELEC member into its own language and notified to the Central Secretariat has the same status as the official versions.

CENELEC members are the national electrotechnical committees of Austria, Belgium, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, Switzerland and United Kingdom.

CENELEC

European Committee for Electrotechnical Standardization
Comité Européen de Normalisation Electrotechnique
Europäisches Komitee für Elektrotechnische Normung

Central Secretariat: rue de Stassart 35, B - 1050 Brussels

Foreword

The text of the International Standard IEC 62356-2:2003, prepared by IEC TC 100, Audio, video and multimedia systems and equipment, was submitted to the formal vote and was approved by CENELEC as EN 62356-2 on 2004-09-01 without any modification.

The following dates were fixed:

- latest date by which the EN has to be implemented
at national level by publication of an identical
national standard or by endorsement (dop) 2005-09-01
- latest date by which the national standards conflicting
with the EN have to be withdrawn (dow) 2007-09-01

Annex ZA has been added by CENELEC.

Endorsement notice

The text of the International Standard IEC 62356-2:2003 was approved by CENELEC as a European Standard without any modification.

CONTENTS

1	Scope	5
2	Normative references	5
3	Introduction	6
4	Encoding	6
4.1	Overview	6
4.2	Pre-processing	8
4.3	Shuffling	12
4.4	Field-frame decision	14
4.5	Discrete Cosine Transform (DCT)	16
4.6	Rate control	17
4.7	Quantization	18
4.8	Entropy coding	18
4.9	Picture data packing	22
4.10	Auxiliary data	28
5	Decoding	31
5.1	Overview	31
5.2	Unpacking	31
5.3	Entropy decoding	31
5.4	Inverse quantization	31
5.5	Inverse DCT	32
5.6	De-shuffling	32
5.7	Post-processing	32
	Annex A (normative) Subsampling filter	33
	Annex B (normative) Channel shuffling	35
	Annex C (normative)	38
	Annex D (normative) VLC tables	41
	Annex ZA (normative) Normative references to international publications with their corresponding European publications	54
	Bibliography	55
	Figure 1 – Encoding block diagram	8
	Figure 2 – Sampling relationships for 1 080/I and 1 080/PsF source and subsampled systems	10
	Figure 3 – Channel division of subsampled 1 080/I and 1 080/PsF signals	11
	Figure 4 – Channel distribution	12
	Figure 5 – Code blocks and basic blocks in channel	13
	Figure 6 – Shuffle block format	13
	Figure 7 – Shuffle block header byte descriptions	14
	Figure 8 – Frame-mode chrominance DCT block reformat	15
	Figure 9 – Field-mode DCT block reformat	16
	Figure 10 – DCT coefficient encoding example	21

Figure 11 – Basic block format	22
Figure 12 – Frame-mode luminance and chrominance cells	22
Figure 13 – Field-mode luminance and chrominance cells	23
Figure 14 – Framemode placement for Offset Mode and Offset Index bits	25
Figure 15 – Fieldmode placement for Offset Mode and Offset Index bits	25
Figure 16 – Packing when quantizer base = 61 or less	26
Figure 17 – Packing when quantizer base = 63	27
Figure 18 – Auxiliary basic block format	28
Figure 19 – Auxiliary data words	30
Figure 20 – Decoding block diagram	31
Figure A.1 – Template for insertion-loss frequency characteristic (Y)	33
Figure A.2 – Passband ripple tolerance (Y)	33
Figure A.3 – Template for insertion-loss frequency characteristic (C_B, C_R)	34
Figure A.4 – Passband ripple tolerance (C_B, C_R)	34
Figure B.1 – 8*8 block segmentation in each channel.....	35
Figure B.2 –Block allocation within a segment.....	36
Table 1 – Data rates associated with source picture rates	6
Table 2 – Definition of signal sampling parameters	9
Table 3 – Data representation	16
Table 4 – DC quantization divisors	18
Table 5 – AC quantization divisors	18
Table 6 – Offset mode and offset index	19
Table 7 – DC coefficient fixed precision	19
Table 8 – Example luminance a.c. coefficient encoding.....	20
Table 9 – Auxiliary basic block data	28
Table 10 – MSB inversion	32
Table B.1 – Equation for TMP1	36
Table B.2 – Values of START_OFFSET for luminance planes	37
Table B.3 – Values of START_OFFSET for chrominance planes	37
Table C.1 – Dynamic range of coefficients	39
Table C.2 – Coefficients for d.c. only transforms	39
Table C.3 – $8_H \times 8_V$ zigzag scan	39
Table C.4 – $4_H \times 8_V$ zigzag scan	40
Table C.5 – $8_H \times 4_V$ zigzag scan	40

VIDEO RECORDING – 12,65 MM TYPE D-11 FORMAT –

Part 2: Picture compression and data stream

1 Scope

This International 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 the 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 $1\,920 \times 1\,080$ pixels and the sampling structures as specified in SMPTE 274M and RP 211 at the following picture rates:

- 24/1,001/PsF;
- 24/PsF;
- 25/PsF;
- 30/1,001/PsF;
- 50/I;
- 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 referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

SMPTE 292M:1998, *Television – Bit-Serial Digital Interface for High-Definition Television Systems*

SMPTE 274M:1998, *Television – $1\,920 \times 1\,080$ Scanning and Analog and Parallel Digital Interfaces for Multiple Picture Rates*

SMPTE RP 211:2000, *Implementation of 24P, 25P and 30P Segmented Frames for $1\,920 \times 1\,080$ Production Format*

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

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

3 Introduction

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

In common with other compression systems, the Type D-11 encoding process uses intra-frame coding (i.e. the coding is bound by the frame period) using the Discrete Cosine Transform (DCT) to provide the data de-correlation 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 $1\,920 \times 1\,080$ pixels is reduced to $1\,440 \times 1\,080$ pixels. For each chrominance channel, the source sampling grid of $960 \times 1\,080$ pixels is reduced to $480 \times 1\,080$ pixels. In the decoder, the output pixel sample grid is restored back to the source format of $1\,920 \times 1\,080$ 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

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

NOTE DCT coding uses a data block size which allows exactly 1 080 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.

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

- 24/1 001 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;
- 30/1 001 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;

- 60/1 001 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 pre-filtered 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 eight samples of eight 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 (15C_B and 15 C_R).

The picture data in each code block shall be compressed by the application of the DCT, 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 1 080 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 nine luminance 8*8 blocks and six 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 5.

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.

NOTE The '**' symbol is used to denote multiplication.

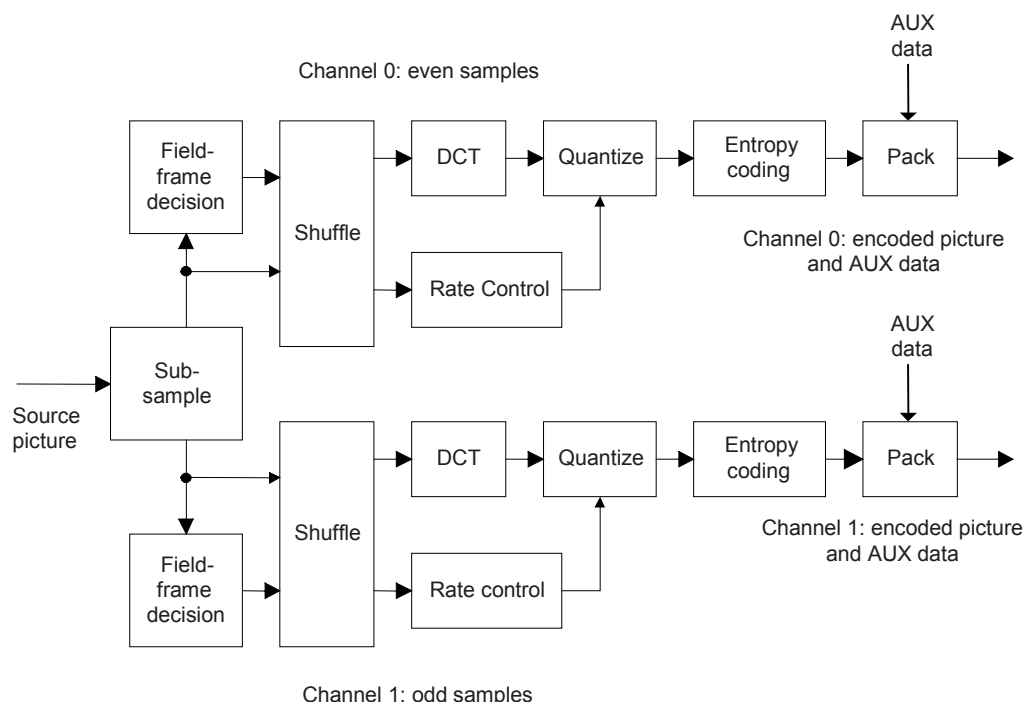


Figure 1 – Encoding block diagram

4.2 Pre-processing

4.2.1 Source picture

The source picture shall be the production aperture as defined in SMPTE 274M having a luminance structure of 1 920*1 080 pixels and a multiplexed chrominance structure of 960*1 080 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.

4.2.2 Vertical sampling process

For 1 080/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 1 080/PsF systems, 1 080 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 2.

4.2.3 Horizontal subsampling process

For the luminance component, all the 1 920 active samples per line shall be subsampled to 1 440 samples per line after a bandwidth limitation filtering process.

For each of the two chrominance components, all the 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 sub-sampled component signals are described in Table 2.

Figure 2 depicts the re-sampled spatial positions of the subsampled components for 1 080/I and 1 080/PsF line-scanning 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 3.

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

For Y samples $r = 0, 1, 2, 3, \dots, 1\,439$

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 3, shall be distributed to channel 0.

Those samples that have ' r ' as an odd number, depicted as a grey circle in Figure 3, shall be distributed to channel 1.

Each luminance (Y) sample channel has rectangular area of 720 samples by 1 080 lines. Each chrominance (C_B , C_R) sample channel has a rectangular area of 240 samples by 1 080 lines respectively as illustrated in Figure 4.

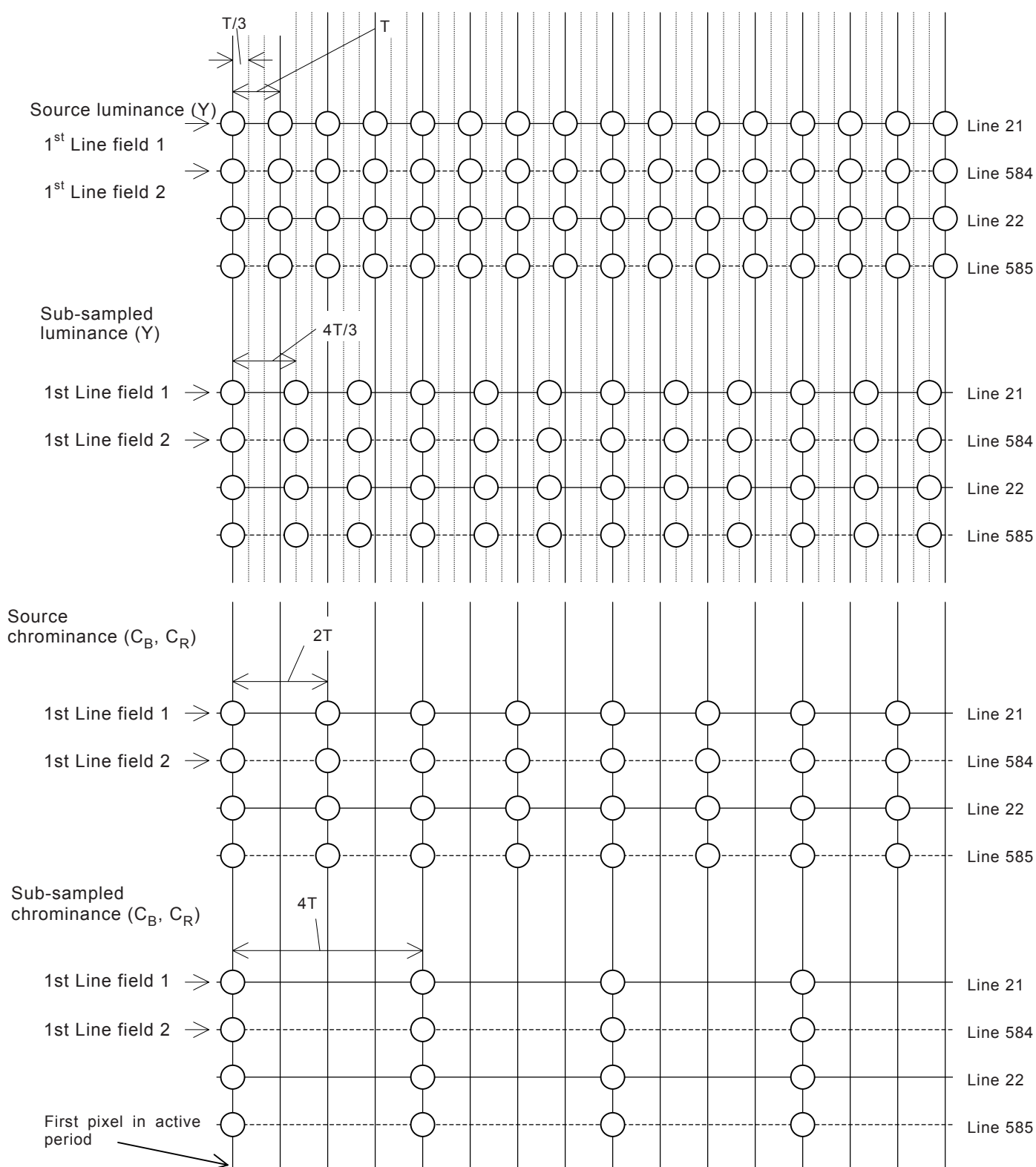
Figure 3 shows the overall structure of the subsampling process.

To avoid alias artifacts, the source format shall be pre-filtered 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.

Table 2 – Definition of signal sampling parameters

Parameter		Source sampling	Subsampling	Channel division
Number of samples per line	Y	1 920	1 440	720
	C_B , C_R	960	480	240
Number of active lines per frame		1 080	1 080	1 080
Quantization		10-bit (0..1 023)	8-bit (0..255)	8-bit (0..255)
Sample levels	Peak range	4 to 1 019	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



NOTE 'T' is the period of the luminance horizontal sampling.

Figure 2 – Sampling relationships for 1 080/I and 1 080/PsF source and subsampled systems

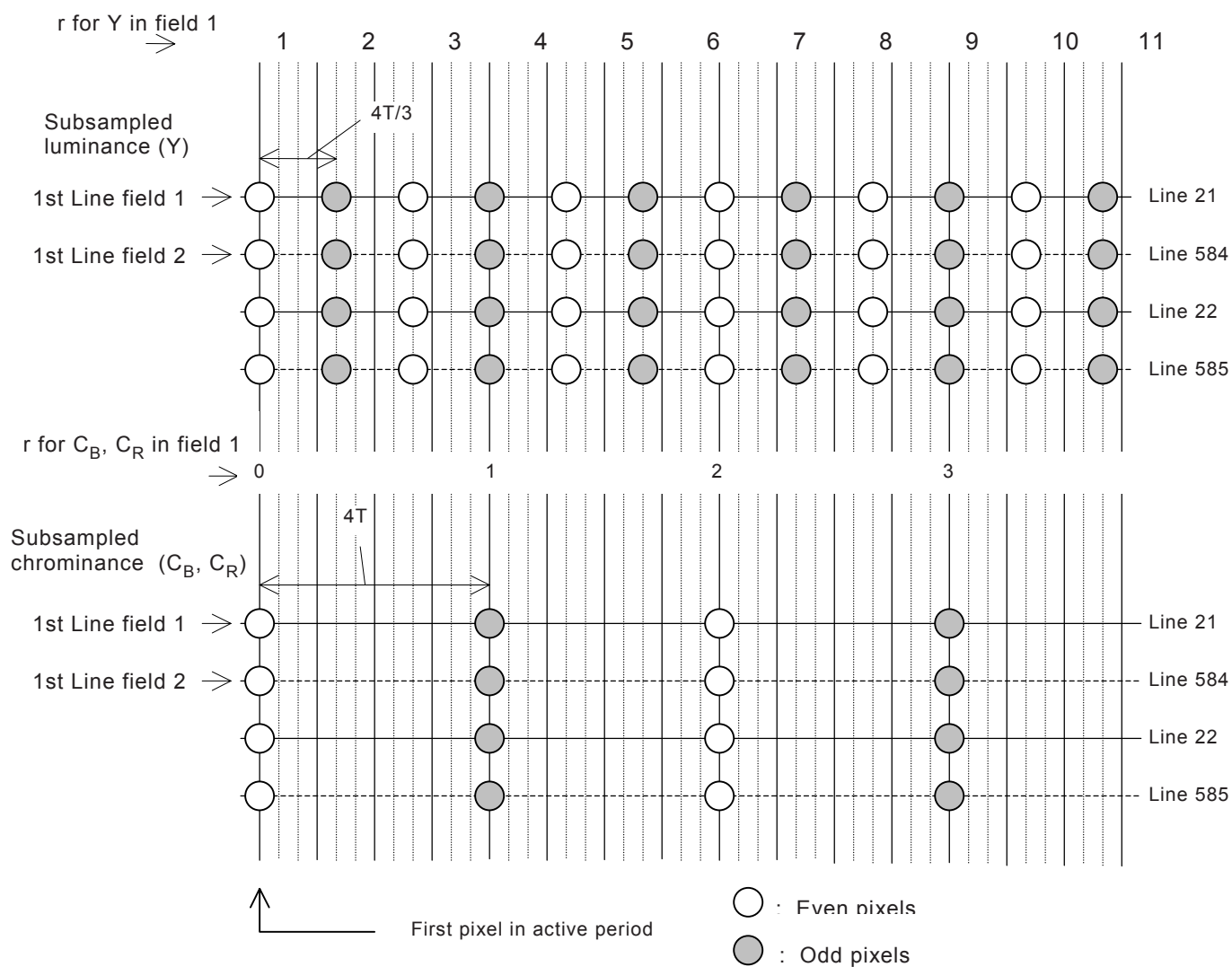


Figure 3 – Channel division of subsampled 1 080/I and 1 080/PsF signals

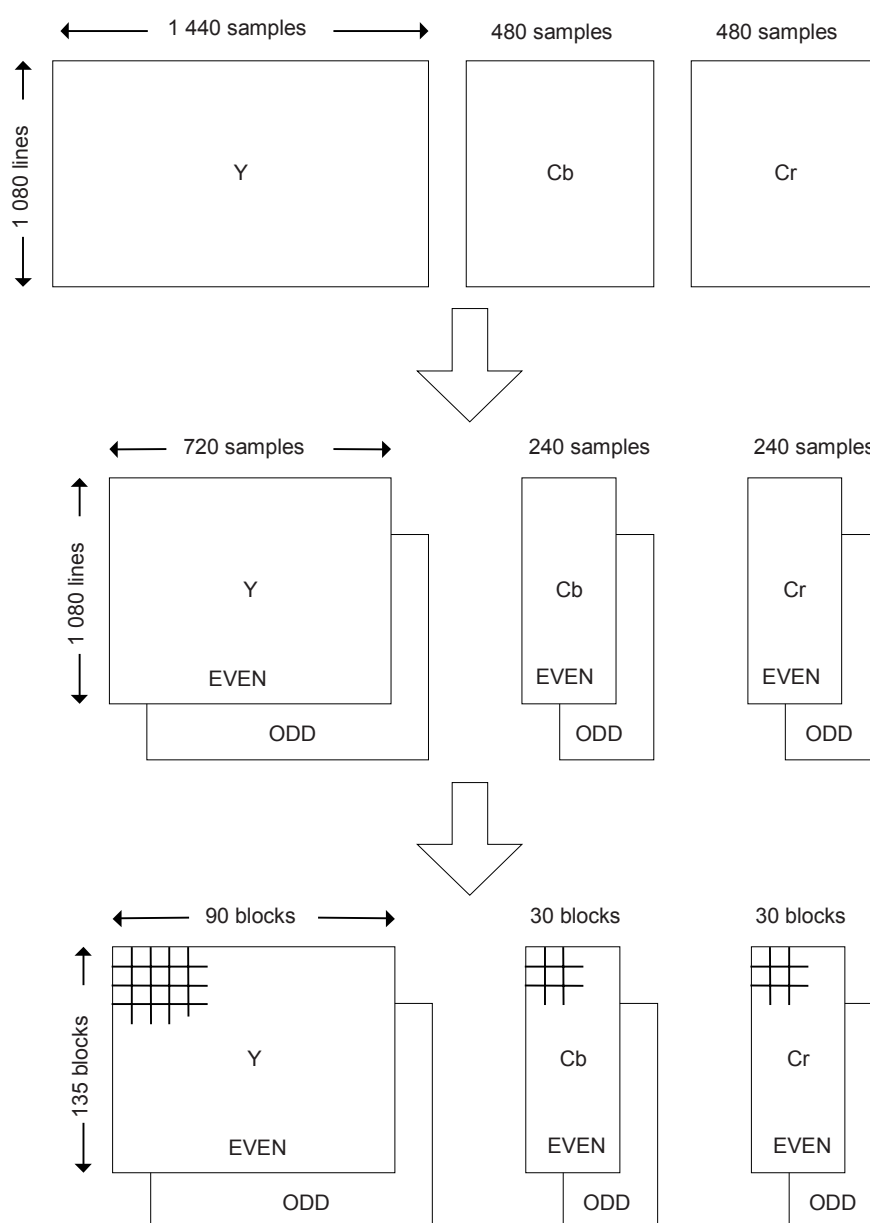


Figure 4 – Channel distribution

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 4. 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 4. The shuffling re-arranges 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 5.

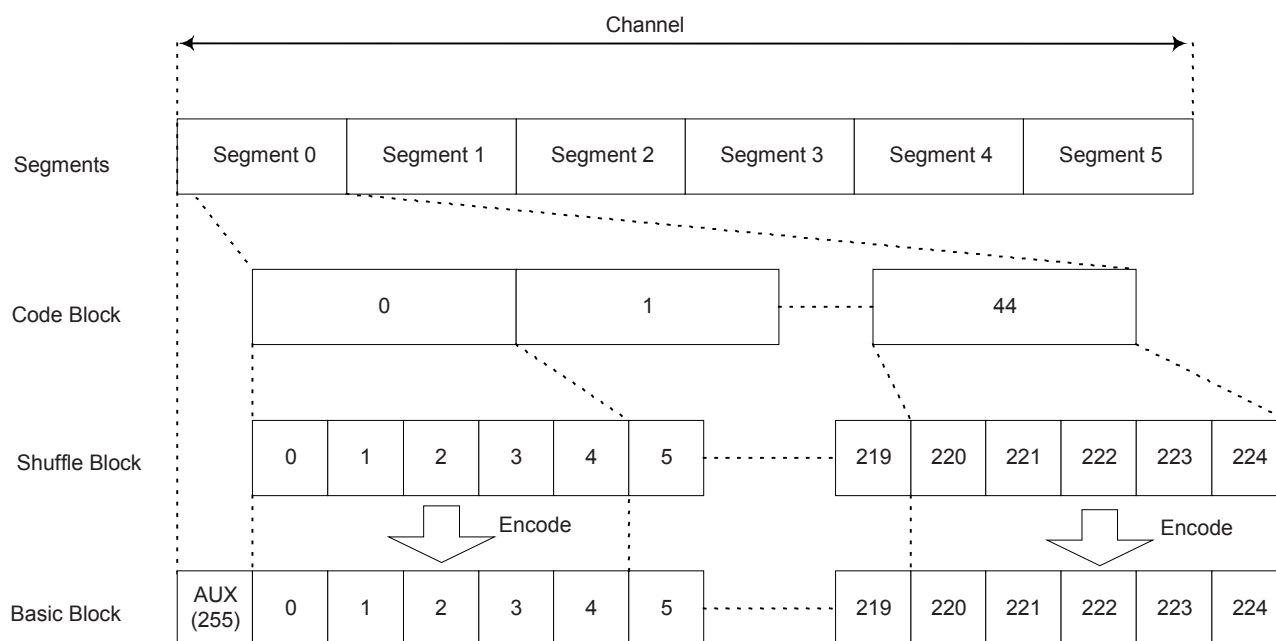


Figure 5 – Code blocks and basic blocks in channel

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

Each shuffle block, defined at the output of the shuffle algorithm, comprises three header bytes, nine luminance 8*8 blocks and six chrominance 8*8 blocks, as shown in Figure 6.

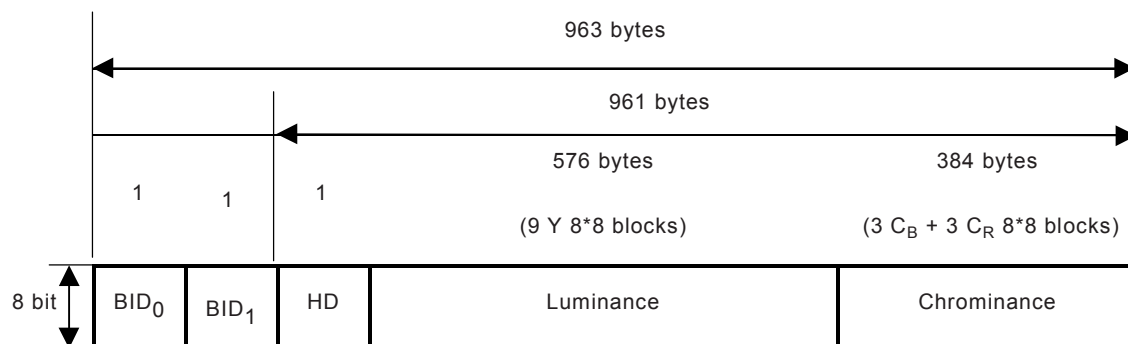


Figure 6 – Shuffle block format

The first header byte, BID₀, shall define the shuffle block number from Figure 5 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 QB_5 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	2nd 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.

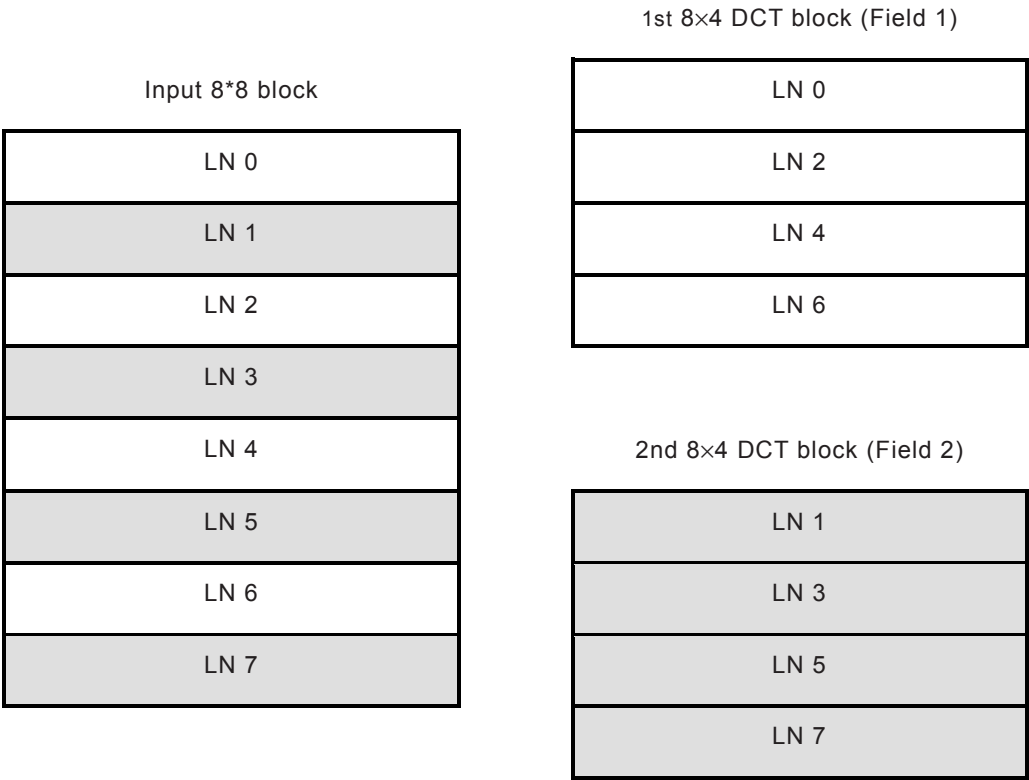


Figure 9 – Field-mode DCT block reformat

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 2's complement form and the chrominance from 2's complement with MSB inversion to 2's 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 d.c. coefficient and a number of a.c. coefficients depending on the DCT block size.