

SMPTE STANDARD

Carriage of Level 65 VC-2 Compressed Video Over the SDTV SDI



Page 1 of 10 pages

Table of Contents	Page
Foreword	2
Intellectual Property	2
Introduction.....	2
1 Scope	3
2 Conformance Notation	3
3 Normative References	4
4 General Description	4
4.1 Source Image.....	4
4.2 Timing Reference Codes	5
4.3 Ancillary Data.....	5
4.4 Audio Data	5
4.5 Time Code.....	5
4.6 Video Payload Identifier	5
4.7 Viewable Content	6
5 VC-2 Level 65 Bitstream Mapping	6
6 Interface Data Format	7
6.1 Compressed Data Mapping	7
6.2 Viewable Picture	8
6.3 Creating a Viewable Picture (Informative)	8
Annex A Effect of Halftone Dither (Informative).....	10

Foreword

SMPTE (the Society of Motion Picture and Television Engineers) is an internationally-recognized standards developing organization. Headquartered and incorporated in the United States of America, SMPTE has members in over 80 countries on six continents. SMPTE's Engineering Documents, including Standards, Recommended Practices and Engineering Guidelines, are prepared by SMPTE's Technology Committees. Participation in these Committees is open to all with a bona fide interest in their work. SMPTE cooperates closely with other standards-developing organizations, including ISO, IEC and ITU.

SMPTE Engineering Documents are drafted in accordance with the rules given in Part XIII of its Administrative Practices.

SMPTE ST 2047-4 was prepared by Technology Committee 32NF.

Intellectual Property

At the time of publication no notice had been received by SMPTE claiming patent rights essential to the implementation of this Standard. However, attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. SMPTE shall not be held responsible for identifying any or all such patent rights.

Introduction

This section is entirely informative and does not form an integral part of this Engineering Document.

This standard defines the provisions for mapping VC-2 bitstreams onto a 270 Mbit/s SDTV SDI link and describes a method by which additional bits are added to ensure that the resultant signal can be viewed on a SDTV SDI monitor, albeit in a somewhat noisy form.

The VC-2 compression standard (SMPTE ST 2042-1) defines a generic intra-frame compression system that can handle all defined video formats and more. A companion recommended practice (SMPTE RP 2047-3) defines constraints (defined as level 65) on the VC-2 specification such that high definition source video images (1080 line interlaced or 720 line progressive) can be compressed such that the resulting compressed bit-rate that can be carried over an SDTV SDI link operating at 270 Mbit/s.

A key feature of this specification is that the transport of compressed data in this way is backwards compatible with installed equipment. The compressed video is mapped entirely into active picture and the signal on the link includes viewable video signal. This aids compatibility with existing equipment such routing and monitoring apparatus and also uncompressed standard definition recording apparatus.

An important feature of this mapping process is that it can be implemented in a practical form using currently available firmware and hardware.

1 Scope

This standard defines the method of mapping VC-2 data onto an SDTV SDI link operating at 270Mbps, based upon the data structures defined in SMPTE ST 259 and ITU-R BT 656 to transport the compressed HD video formats listed in Table 1. The compressed HD video formats supported by this standard are those supported by the low delay profile of VC-2 at level 65. These formats are 1080I at 50 or 60/1.001 fields/s (i.e., SMPTE ST 274 systems 5 and 6), and 720P at 50 or 60/1.001 frames/s (i.e., SMPTE ST 296 systems 2 and 3). Additionally VC-2 level 65 supports the compression of SMPTE ST 274 systems 5 and 6 in which the horizontal sample rate has been reduced to 1440 pixels.

Table 1 – Supported HD Video Formats

Format Name	Frame Height	Frame Width	Frame Rate (Hz)	Scanning	Format Standard	System Number
1920x1080/60/I	1080	1920	29.97	interlace	SMPTE ST 274	5
1920x1080/50/I	1080	1920	25	interlace	SMPTE ST 274	6
1440x1080/60/I	1080	1440	29.97	interlace	SMPTE RP 2047-3	5
1440x1080/50/I	1080	1440	25	interlace	SMPTE RP 2047-3	6
1280x720/60/P	720	1280	59.94	progressive	SMPTE ST 296	2
1280x720/50/P	720	1280	50	progressive	SMPTE ST 296	3

The SDTV SDI link defined herein supports the carriage of embedded audio, ancillary data and payload identification.

A signal coded in accordance with this standard incorporates a noisy but recognizable picture that can be viewed using a standard SMPTE ST 259 / ITU-R BT 656 video monitor.

2 Conformance Notation

Normative text is text that describes elements of the design that are indispensable or contains the conformance language keywords: "shall", "should", or "may". Informative text is text that is potentially helpful to the user, but not indispensable, and can be removed, changed, or added editorially without affecting interoperability. Informative text does not contain any conformance keywords.

All text in this document is, by default, normative, except: the Introduction, any section explicitly labeled as "Informative" or individual paragraphs that start with "Note:"

The keywords "shall" and "shall not" indicate requirements strictly to be followed in order to conform to the document and from which no deviation is permitted.

The keywords, "should" and "should not" indicate that, among several possibilities, one is recommended as particularly suitable, without mentioning or excluding others; or that a certain course of action is preferred but not necessarily required; or that (in the negative form) a certain possibility or course of action is deprecated but not prohibited.

The keywords "may" and "need not" indicate courses of action permissible within the limits of the document.

The keyword “reserved” indicates a provision that is not defined at this time, shall not be used, and may be defined in the future. The keyword “forbidden” indicates “reserved” and in addition indicates that the provision will never be defined in the future.

A conformant implementation according to this document is one that includes all mandatory provisions (“shall”) and, if implemented, all recommended provisions (“should”) as described. A conformant implementation need not implement optional provisions (“may”) and need not implement them as described.

Unless otherwise specified, the order of precedence of the types of normative information in this document shall be as follows: Normative prose shall be the authoritative definition; Tables shall be next; followed by formal languages; then figures; and then any other language forms.

3 Normative References

Note: All references in this document to other SMPTE documents use the current numbering style (e.g. SMPTE ST 12-2:2008) although, during a transitional phase, the document as published (printed or PDF) may bear an older designation (such as SMPTE 12M-2-2008). Documents with the same root number (e.g. 12-2) and publication year (e.g. 2008) are functionally identical.

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 ST 12-2:2008, Television — Transmission of Time Code in the Ancillary Data Space

SMPTE ST 259:2008, Television — SDTV Digital Signal/Data — Serial Digital Interface

SMPTE ST 272:2004, Television — Formatting AES/EBU Audio and Auxiliary Data into Digital Video Ancillary Data Space

SMPTE ST 274:2008, Television — 1920 x 1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequences for Multiple Picture Rates

SMPTE ST 291:2010, Ancillary Data Packet and Space Formatting

SMPTE ST 296:2001, Television — 1280 x 720 Progressive Image Sample Structure — Analog and Digital Representation and Analog Interface

SMPTE ST 352:2011, Video Payload Identification Codes for Serial Digital Interfaces

SMPTE ST 2042-1:2009, VC-2 Video Compression

SMPTE RP 2047-3:2011, VC-2 Level 65 Compression of High Definition Video Sources for use with a Standard Definition Infrastructure

Recommendation ITU-R BT.656-5 (12/07), Interface for Digital Component Video Signals in 525-Line and 625-Line Television Systems Operating at the 4:2:2 Level of Recommendation ITU-R BT.601

4 General Description

4.1 Source image formats

The data for this interface shall be derived from the following source image sampling formats as specified in SMPTE RP 2047-3:

- SMPTE ST 274 systems 5 and 6; i.e., 1920 pixels by 1080 lines, interlaced, 10 bit, 4:2:2 video at either 50 field/s or 60/1.001 field/s.

- SMPTE ST 274 systems 5 and 6, in which the horizontal sample count has been reduced to 1440 pixels.
- SMPTE ST 296 systems 2 and 3; i.e., 1280 pixels by 720 lines, progressive, 10 bit, 4:2:2 video at either 50 frames/s or 60/1.001 frames /s.

4.2 Timing Reference Codes

Timing Reference codes shall be in conformance with those specified in either SMPTE ST 259 or ITU-R BT.656-5 corresponding to the field or the frame rate of the source image format.

For interlaced source image formats their field rate shall be the same as field rate on the SDI link.

For progressive source image formats their frame rate shall be the same as the field rate on the SDI link.

4.3 Ancillary Data

Ancillary data, if present, shall be mapped into the blanking area in conformance with SMPTE ST 291.

4.4 Audio Data

Audio data, if present, shall be mapped into the ancillary data space in conformance with SMPTE ST 272.

4.5 Time Code

Time code, if present, shall be mapped into the ancillary data space in conformance with SMPTE ST 12-2.

4.6 Video Payload Identifier

The payload identifier shall be present and shall be mapped into the ancillary data space in conformance with SMPTE ST 352.

The first byte of the payload identifier shall be used to identify the combination of video payload format and digital interface transport. For an interface in accordance with this standard Byte 1 shall be set to B2h.

Bytes 2 and 3 of the payload identifier shall be set in accordance with the picture rate, sampling structure, and aspect ratio etc. of the image being carried on the interface as shown in Table 2.

Byte 4 of the payload identifier shall be used to identify the compressed payload. The meaning of byte 4 shall depend on the values of bytes 2 and 3 as shown in Table 3.

Table 2 – Video Payload Identifier Definition

Bits	Byte 1	Byte 2	Byte 3	Byte 4
Bit 7	1	0 (Interlaced transport)	1 (Image aspect ratio: 16:9)	0
Bit 6	0	0 (Interlaced picture)	0 (Reserved)	0
Bit 5	1	00 (VC-2 compression)	0 (Reserved)	0
Bit 4	1		0 (Reserved)	0
Bit 3	0	Picture Rate (see SMPTE ST 352 Table 2)	0h (Sampling structure: 4:2:2 Y/Cb/Cr)	0
Bit 2	0			See Table 3
Bit 1	1			See Table 3
Bit 0	0			See Table 3

Table 3 – Compressed Payload Formats

Byte 2	Byte 3	Byte 4	Format Name
0x06 (Picture rate = 30/1.001 Hz)	0x80	0x01	1920x1080/60/I
0x05 (Picture rate = 25 Hz)	0x80	0x02	1920x1080/50/I
0x06 (Picture rate = 30/1.001 Hz)	0x80	0x05	1440x1080/60/I
0x05 (Picture rate = 25 Hz)	0x80	0x06	1440x1080/50/I
0x0A (Picture rate = 60/1.001 Hz)	0x80	0x03	1280x720/60/P
0x09 (Picture rate = 50 Hz)	0x80	0x04	1280x720/50/P

Byte 2, bits 4 and 5 shall be set to 00 to indicate that VC-2 syntax is used to compress the source video signal.

Byte 4 shall be set as shown in Table 3 to indicate the carriage of the VC-2 compressed payload and the viewable video data according to this standard.

The formats indicated in Table 3 are per SMPTE RP 2047-3.

4.7 Viewable Content

A signal coded in accordance with this interface should produce a recognizable version of the compressed payload (although noisy and coarsely quantized) when decoded and displayed by a SMPTE ST 259 / ITU-R BT 656 compliant receiving interface as if it were a standard definition interlaced signal.

5 VC-2 Level 65 Bitstream Mapping

This section defines which data from a VC-2 Level 65 bitstream according to SMPTE RP 2047-3 is mapped into the SDI signal.

Consecutive VC-2 pictures shall be mapped to consecutive SDI fields. Hence each SDI frame is used to transport two VC-2 compressed pictures of the source image format.

Note: Compliance with VC-2 level 65 ensures that VC-2 pictures correspond to either a field or a frame in the input source video depending on whether the input source video is interlaced or progressive respectively.

A compressed VC-2 stream comprises data and metadata. The metadata describes the parameters of the coded signal and the way in which it is encoded. The data describes the details of the specific video signal that has been compressed. When a VC-2 stream is transported over a SDI link according to this specification, the stream metadata shall be inferred from the payload identifier (defined in Section 4.6). Therefore, it is only the data from the VC-2 stream that is mapped to the SDI link.

The signal data part of a VC-2 stream shall comprise the transform data part, as defined in Sections 12.4 and 12.5 of SMPTE ST 2042-1, VC-2 Video Compression. The signal data part of each VC-2 picture shall be

mapped into the active picture area of one field within the SDI signal. In this specification an active field means 720 pixels by either 288 or 243 standard definition lines for 50 or 59.94 field/s video respectively.

No metadata from the VC-2 stream shall be mapped into the SDI signal (except by inference from the payload identifier).

The complete VC-2 stream (with the exception of VC-2 "Picture Numbers" defined in Section 11.1 of the VC-2 specification) can be reconstructed from the transform data present on the SDI and the metadata inferred from the payload identifier.

Note: The metadata to be inferred from the payload identifier is defined and enumerated in detail in SMPTE RP 2047-3, VC-2 Level 65 Compression of High Definition Video Sources for use with Standard Definition Infrastructure.

VC-2 "Picture Numbers" provide unique identifiers for pictures within a sequence. Picture numbers within a VC-2 sequence are required to increment by one for each successive picture within a given sequence. Between sequences they are unconstrained and so picture numbers, starting from any arbitrary value, may be inserted in a decoded VC-2 stream while retaining compliance with the specification.

Decoding apparatus may decode the VC-2 stream implied as in the previous paragraph to produce an uncompressed video signal.

Note: The coding parameters defined in SMPTE RP 2047-3 ensure that the compressed transform data, plus the bits added for the viewable picture, occupy precisely the space allocated within the SDI signal for the active picture.

6 Interface Data Format

6.1 Compressed Data Mapping

The VC-2 compressed video data shall be transported in the least-significant 8 bits of the (10-bit) video words as shown in Table 4. The first video word shall contain the first byte of the VC-2 transform data. Subsequent bytes of the transform data shall be transported in the 8 least-significant bits of consecutive video words, irrespective of whether the video words correspond to Cb, Y or Cr data. No VC-2 compressed video data shall be transported other than in the least-significant 8 bits of video words.

Table 4 – Mapping of compressed video data to interface words

Word	Bit number									
	9 (MSB)	8	7	6	5	4	3	2	1	0 (LSB)
XYZ (C)	1	F	0	0	P3	P2	P1	P0	0	0
XYZ (Y)	1	F	0	0	P3	P2	P1	P0	0	0
Cb data			Byte 0 of VC-2 Transform Data							
Y data			Byte 1 of VC-2 Transform Data							
Cr data			Byte 2 of VC-2 Transform Data							
Y data			Byte 3 of VC-2 Transform Data							
Cb data			Byte 4 of VC-2 Transform Data							
Y data			Byte 5 of VC-2 Transform Data							
Cr data			Byte 6 of VC-2 Transform Data							
Y data			Byte 7 of VC-2 Transform Data							
•			•							
•			•							
•			•							

(MSB – Most Significant Bit, LSB – Least Significant Bit, MS – Most Significant, LS – Least Significant)

The shaded region, comprising the 2 most-significant bits of each video word in the active picture area, contains the viewable picture information as described in Section 0.

6.2 Viewable Picture

The viewable picture shall represent the contents of the compressed video. The two most-significant bits of the 10-bit interface words shall be used to carry a viewable version of the input picture, so that if the SDI signal is viewed as if it were standard SDI video, the content will be a recognizable, although noisy and coarsely quantized, video. This makes it possible to view the signal for the purpose of identification and to give confidence that the coder is working.

The encoding method used to generate the viewable image in the top 2 most-significant bits of each video word is not defined in this specification although guidance is given in Section 6.3.

Whatever encoding method is used to generate the viewable picture, the video words shall not contain the prohibited synchronization codes as specified in SMPTE ST 259 (i.e., 000_h to 003_h and 3FC_h to 3FF_h).

6.3 Creating a Viewable Picture (Informative)

Generating the viewable picture can be achieved as follows:

- a) A version of the source video that has been compressed is generated corresponding to the format of the SDI signal. This is the unquantized viewable signal. In order to generate this signal the HD signal must be down-converted to the SD raster. For example 1920 pixels would need to be reduced to only 720 pixels. A simple way to do this is to extract 3 output pixels from every 8 input pixels (720/1920 = 3/8).

Selecting the required output pixel can be accomplished by incrementing a “phase accumulator” by $8/3$ for each output pixel and rounding the result to yield the address of the required input pixel. Similarly the 540 lines of an HD field can be reduced to the 243 lines of a (59.94field/s) SD field by selecting input lines on the basis of a phase accumulator incremented by $540/243$ for each output line. Alternatively more sophisticated linear filtering techniques can be used to effect the down-conversion, but such techniques yield only a small improvement in the quality of the viewable picture.

- b) A video signal, corresponding to the compressed data formatted as in the SDI signal with the two most-significant bits set to zero, is subtracted from the unquantized viewable signal. A two dimensional halftone dither signal is added to this difference signal. The result must be limited to a valid 10-bit range. The dither consists of a repeating 8x8 pattern as follows:

Table 5 – 2D halftone dither

0	128	32	160	8	136	40	168
192	64	224	96	200	72	232	104
48	176	16	144	56	184	24	152
240	112	208	80	248	120	216	88
12	140	44	172	4	132	36	164
204	76	236	108	196	68	228	100
60	188	28	156	52	180	20	148
252	124	220	92	244	116	212	84

The dither is added independently to the luminance, Cb color-difference and Cr color-difference components, where the row selected is determined by the line number evaluated modulo-8, and the column selected is determined by the pixel number (e.g., 0-1919 for luminance, 0-959 for color-difference) evaluated modulo-8.

- c) The least-significant 8-bits of the resulting video are discarded and replaced by the 8-bit values of the compressed data.
- d) If the final 10-bit value corresponds to a prohibited synchronization code as specified in SMPTE ST 259 (000_h to 003_h and 3FC_h to 3FF_h), the most-significant two bits shall be modified as follows:

00 000000 xx is modified to 01 000000 xx
 11 111111 yy is modified to 10 111111 yy

Annex A Effect of Halftone Dither (Informative)

The effect of the halftone dither can be seen from the following illustrations (magnified so the individual pixels are visible):

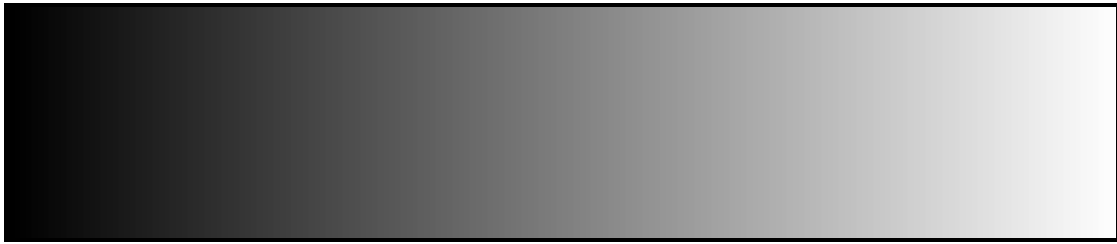


Figure A.1 – Original linear ramp

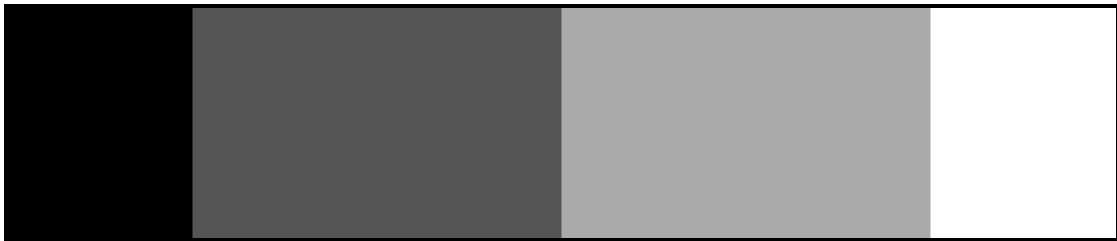


Figure A.2 – Ramp quantized to two bits

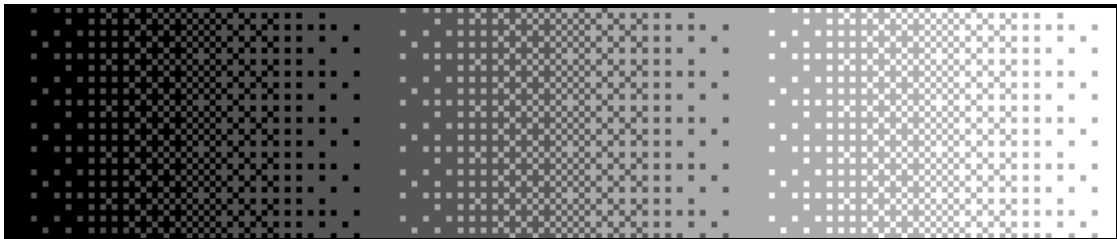


Figure A.3 – Ramp quantized to two bits with 2D halftone dither

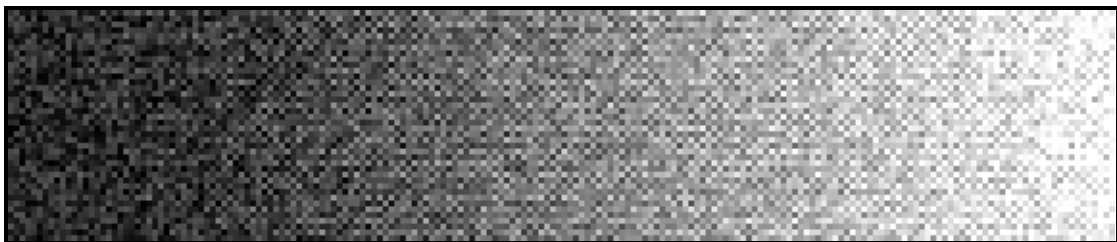


Figure A.4 – Quantized ramp with dither and random data in LSBs