SMPTE ST 2110-20:2022
Revision of SMPTE ST 2110-20:2017

SMPTE STANDARD

Professional Media Over Managed IP Networks: Uncompressed Active Video

Table of Contents Page

Approved December 14, 2022

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 its Standards Operations Manual. This SMPTE Engineering Document was prepared by Technology Committee 32NF.

This revision extends the 2017 original publication, updating normative references to current versions thereof, and improving certain normative language based on the related PICS document. Some provisions from 2110-21:2017 are now included in this document. This revision also adds definitions and support for Key (alpha) transport and adds signaling support for the SMPTE ST 2115 LOGS3 transfer characteristic.

A previously published version of this document with the approval date of 2022-03-28 contained typos in the dates of certain normative references pointing at non-existent documents. These typos have been corrected in this version (approval date 2022-12-14).

Intellectual Property

At the time of publication no notice had been received by SMPTE claiming patent rights essential to the implementation of this Engineering Document. 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.

The capability and capacity of IP networking equipment has improved steadily, enabling the use of IP switching and routing technology to transport and switch video, audio, and metadata essence within television facilities. Existing standards such as SMPTE ST 2022-6 have gained use in this application, but there is a desire in the industry to switch different essence elements separately.

This family of SMPTE engineering documents builds on the work of Video Services Forum (VSF) Technical Recommendations TR03 and TR04, and on AES67, documenting a system for transporting various essence streams over IP networks. The system is designed to be extensible to a variety of essence types.

SMPTE ST 2110-10 covers the system as a whole, the timing model, and common requirements across all essence types.

This standard documents the transport of uncompressed active video in such systems, using an RTP format based on IETF RFC 4175.

1 Scope

This standard specifies the real-time, RTP-based transport of uncompressed active video essence over IP networks. An SDP-based signaling method is defined for image technical metadata necessary to receive and interpret the stream.

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; then formal languages; then figures; and then any other language forms.

3 Normative References

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

IETF RFC 3550 Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, DOI 10.17487/RFC3550, July 2003, https://www.rfc-editor.org/info/rfc3550

IETF RFC 4566 Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", DOI 10.17487/RFC4566, July 2006, https://www.rfc-editor.org/info/rfc4566

IETF RFC 8285 Singer, D., Desineni, H., and R. Even, Ed., "A General Mechanism for RTP Header Extensions", DOI 10.17487/RFC8285, October 2017, https://www.rfc-editor.org/info/rfc8285

ISO 11664-1:2007 Colorimetry -- Part 1: CIE standard colorimetric observers

Recommendation ITU-R BT.601-7 (03/2011) Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios

Recommendation ITU-R BT.709-6 (06/2015) Parameter values for the HDTV standards for production and international programme exchange

Recommendation ITU-R BT.1886 (03/2011) Reference electro-optical transfer function for flat panel displays used in HDTV studio production

Recommendation ITU-R BT.2020-2 (10/2015) Parameter values for ultra-high definition television systems for production and international programme exchange

Recommendation ITU-R BT.2100-2 (07/2018) Image Parameter Values for High Dynamic Range Television for use in Production and International Programme Exchange

SMPTE RP 157:2012 Key and Alpha Signals

SMPTE ST 428-1:2006 D-Cinema Distribution Master — Image Characteristics

SMPTE ST 2065-1:2012 Academy Color Encoding Specification (ACES)

SMPTE ST 2065-3:2012 Academy Density Exchange Encoding (ADX) — Encoding Academy Printing Density (APD) Values

SMPTE RP 2077:2013 Full-Range Image Mapping

SMPTE ST 2110-10:2022 Professional Media over Managed IP Networks: System Timing and Definitions

SMPTE ST 2110-21:2022 Professional Media over Managed IP Networks: Traffic Shaping and Delivery Timing for Uncompressed Active Video

SMPTE ST 2115:2019 Free Scale Gamut and Free Scale Log Characteristics of Camera Signals

4 Terms and Definitions

For the purposes of this document, the terms and definitions of SMPTE ST 2110-10 and the following apply.

4.1 Sample Row Number

numerical indication of the position of a row of image samples within the sample array corresponding to the visible (active) picture area

4.2 Sample Row

horizontal collection of image samples spanning the entire width of the image

5 Textual Conventions

5.1 SDP Parameters and Values

The names and values of SDP Media Type parameters within the text of this document are formatted using a monospaced font (such as Courier) except when they appear in section headings.

6 Uncompressed Active Video RTP Essence Format

6.1 Payload Design and Packet format

6.1.1 General Provisions

The sample arrays of the active video essence shall be transported using RTP (IETF RFC 3550), subject to the constraints and payload definition below.

Unless otherwise noted, multi-octet fields within the RTP Header, RTP Payload Header, and RTP Payload shall be transmitted in Network Byte Order (most significant byte first). When represented in bit-field diagrams such as Figure 1 or Figure 2, the most significant bits of multi-bit fields shall occupy the lowest-numbered bit index positions (left-most positions in the figures), and shall be transmitted first.

The image technical metadata necessary to receive and interpret the RTP stream shall be communicated via SDP as defined in section 7.

Traffic shaping and transmission timing of the RTP stream shall be in accordance with the Network Compatibility Model compliance definitions specified in SMPTE ST 2110-21 for Narrow Senders (Type N), Narrow Linear Senders (Type NL), or Wide Senders (Type W).

Where the Media Clock is locked to the timestamping reference clock, traffic shaping and delivery timing shall also be in accordance with the Virtual Receiver Buffer Model compliance definitions specified in SMPTE ST 2110-21 for Narrow Senders (Type N), Narrow Linear Senders (Type NL), or Wide Senders (Type W).

Receivers shall conform to one of the types specified in ST 2110-21:2022 section 7.2.

6.1.2 RTP Header

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| V |P|X|  CC   |M|     PT      |       Sequence Number         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Time Stamp                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              SSRC                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1 — RTP Header

The RTP Packet Header is illustrated in Figure 1.

The fields of the RTP packet header and their order shall be as defined in IETF RFC 3550. The following additional constraints shall apply:

Payload Type (PT): 7 bits — The Payload Type field shall refer to the dynamically allocated payload type as specified in SMPTE ST 2110-10 section 6.2 “Real-Time Transport Protocol (RTP)”.

Timestamp: 32 bits — The Timestamp field shall contain the RTP Timestamp as specified in SMPTE ST 2110-10.

SSRC: 32 bits — The SSRC field shall be as specified in IETF RFC 3550

Marker bit (M): 1 bit — For progressive scan video, the marker bit shall be set to 1 to denote when this RTP packet is the last packet carrying video essence data for a video frame, and set to 0 for all other packets.

For interlaced video, the marker bit shall be set to 1 to denote when this RTP packet is the last packet carrying video essence data for a video field, and set to 0 for all other packets.

Sequence Number: 16 bits — The RTP header sequence number field shall contain the 16 low order bits of the extended 32-bit RTP packet sequence counter.

Extension bit (X): 1 bit — When this bit is set, an RTP header extension is present immediately following the SSRC field. If present, the header extension shall be compliant to IETF RFC 8285.

6.1.3 Media Clock, RTP Clock, and RTP Timestamps

The Media Clock, RTP Clock, and RTP Timestamps shall comply with the provisions of SMPTE ST 2110-10. The RTP Clock rate for streams compliant to this standard shall be 90 kHz.

All RTP packets which are part of the same progressive frame shall contain the same RTP Timestamp value. All RTP packets which are part of the same interlaced field shall contain the same RTP Timestamp value.

6.1.4 RTP Payload Header – Extended Sequence Number and Sample Row Data Headers

An RTP Payload Header containing two Sample Row Data Headers is shown in Figure 2.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Extended Sequence Number   |           SRD Length          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|F|     SRD Row Number          |C|         SRD Offset          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       SRD Length              |F|        SRD Row Number       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|C|     SRD Offset              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2 — RTP Payload Header with Extended Sequence Number and Sample Row Data Headers

The RTP Payload Header shall include the Extended Sequence Number, followed by one, two, or three Sample Row Data (SRD) Headers.

Extended Sequence Number: 16 bits — The Extended Sequence Number field shall contain the 16 high order bits of the extended 32-bit sequence number.

Each Sample Row Data Header shall consist of:

SRD Length: 16 bits — The SRD Length field shall indicate the number of octets of data included from the indicated sample row, and shall be a multiple of the pgroup octet length. An SRD length of zero shall be forbidden except in the case where there is exactly one SRD Header – in which case the SRD length of zero indicates that no sample row data follows this header.

Field Identification (F): 1 bit — The Field Identification bit shall be set to 0 to signal data in the (temporally) first field and set to 1 to signal data in the second field.

For progressive scan data the Field Identification bit shall be set to zero, except in the case of Progressive segmented Frame (PsF) data, in which case the F bit indicates the segment.

SRD Row Number: 15 bits — The SRD Row Number field shall refer to Sample Row Numbers within the Sample Array, and shall start at 0 at the top of the transmitted image.

In the case of interlaced images, each field shall start at 0 at the top of the field.

In the case of Progressive segmented Frame (PsF) data, the sample row numbering shall start at 0 at the top of each segment.

SRD Row Number shall only increase within the field or frame (sample rows are sent in order from top to bottom).

Continuation (C): 1 bit	The Continuation bit shall be set to 1 if an additional Sample Row Data Header follows the current Sample Row Data Header in the RTP Payload Header, which signals that the RTP packet is carrying data for more than one sample row. The Continuation bit shall be set to 0 otherwise.
SRD Offset: 15 bits	The SRD Offset field shall contain the horizontal position, within the full-bandwidth image pixel matrix, of the first full-bandwidth sample in the associated Sample Row Data Segment, where a value of 0 corresponds to the pixel at the left edge of the image.

6.1.5 Additional constraints on the payload format

Successive RTP packets may contain data from parts of the same sample row (with an incremented RTP sequence number, but the same timestamp), if it is necessary to fragment a line.

For 4:2:0 Sampling System with progressive video, the SRD Row Number shall be set to the first Sample Row of the pair of grouped sample rows. Only every other Sample Row shall be signaled in the SRD Row Number field.

SRD Offset shall only increase within the same sample row (samples are sent in order from left to right across the row).

The fields of an interlaced image shall be transmitted in time order, first field first.

For interlaced systems and Progressive segmented Frame (PsF) systems, if the `height` is even, then the lines shall be divided evenly between the fields (or segments), and if the `height` is odd, the temporally first field (or segment) shall contain one more line than the temporally second field.

In progressive systems, a single RTP packet shall not contain samples from more than one frame. In interlaced systems, a single RTP packet shall not contain samples from more than one field. In Progressive segmented Frame (PsF) systems, a single RTP packet shall not contain samples from more than one segment.

In interlaced systems, for the purposes of image reconstruction, the sample rows of the temporally second field shall be displaced vertically “below” the like-numbered sample rows of the temporally first field. Similarly, in Progressive segmented Frame (PsF) systems, the sample rows of the second segment shall correspond to spatial locations below the like-numbered sample rows of the first segment.

Note: Devices that provide a gateway between this standard and SDI ought to be mindful of the distinction between these zero-based Sample Row Numbers in the SRD Line Number field, and the corresponding location within the SDI Interface when interpreting or creating SDI.

6.2 RTP Payload: Sample Row Data Segments

6.2.1 General Provisions

The Sample Row Data Headers shall be followed by Sample Row Data Segments, except in the special case where there is exactly one Sample Row Data Header and it indicates an SRD Length of zero. The order of the Sample Row Data Segments in the RTP payload shall correspond to the order of the Sample Row Data Headers that precede them. For each Sample Row Data Header, there shall be a Sample Row Data Segment in the payload (with the exception of the special case noted above).

RTP packets at the end of the field or frame may contain padding octets after the last Sample Row Data Segment, for the purposes noted in sections 6.3.2 and 6.3.3.

RTP Packets shall not contain more than three Sample Row Data Headers.

Each Sample Row Data Segment shall include data from either a complete sample row, or a portion of a sample row. The length of each Sample Row Data Segment shall in all cases be an integer multiple of the pgroup size in octets. The UDP size of each RTP packet shall not be larger than the prevailing UDP Size Limit (Standard or Extended) as defined in SMPTE ST 2110-10.

It is possible that the Sample Row length is not evenly divisible by the number of samples in a pgroup, causing the final pgroup of the sample row to be only partially filled with sample data. In such situations, the sender shall fill the remaining sample positions of the final pgroup with zero and the receiver shall ignore the fill data. The Length value in the Sample Row Data Header shall include this fill data.

6.2.2 pgroup Size and Construction

A pgroup is the minimal group of samples that align to an octet boundary. Every pgroup shall consist of an integer number of octets. Pgroups shall not be fragmented across packets and shall not represent samples from more than one image source array line (or two source array lines in the case of 4:2:0 sampling).

The definition of which samples are represented in the pgroup shall be determined by the SDP parameters (as detailed in section 7). The number, position, and ordering of the samples within the pgroup is determined by the mandatory `sampling` metadata item defined in section 7.2. If color sub-sampling is used, sub-samples shall only be shared within the pgroup.

'pgroup coverage' in the construction tables below refers to a contiguous portion of the sample array, possibly spanning adjacent lines within a field or frame.

If only one sample of a given component is present in a pgroup, it is un-numbered in the Sample Order column. In the cases where more than one sample of the same component is present within a pgroup, numerical indices ($C0'_B$, $C1'_B$, $C2'_B$, etc.) are used in order to differentiate the samples – the lowest numbered index being left-most within the image.

The sizes and constructions of pgroups are defined in the following sections.

6.2.3 pgroup construction – 4:4:4 Sampling System

When using the 4:4:4 sampling system, pgroups shall be constructed as described in Table 1.

Table 1 — Construction of 4:4:4 pgroups

sampling	depth	pgroup size (octets)	pgroup coverage (pixels)	Sample Order
`YCbCr-4:4:4` `CLYCbCr-4:4:4`	8	3	1	C'_B,Y’,C’R
	10	15	4	$C0'_B$,Y0’,C0’R, C1’B,Y1’,C1’R ,C2' B, Y2’,C2’R ,C3’B,Y3’,C3’R
	12	9	2	$C0'_B$,Y0’,C0’R ,C1’B,Y1’,C1’R
	16, 16f	6	1	C'_B,Y’,C’R
`ICtCp-4:4:4`	8	3	1	C_T,I,CP
	10	15	4	$C0_T$,I0,C0P, C1T,I1,C1P ,C2 T, I2,C2P ,C3T,I3,C3P
	12	9	2	$C0_T$,I0,C0P ,C1T,I1,C1P
	16, 16f	6	1	C_T,I,CP
`RGB` *(linear)*	8	3	1	R, G, B
	10	15	4	R0, G0, B0, R1, G1, B1, R2, G2, B2, R3, G3, B3
	12	9	2	R0, G0, B0, R1, G1, B1
	16, 16f	6	1	R, G, B
`RGB` *(non-linear)*	8	3	1	R’, G’, B’
	10	15	4	R0’, G0’, B0’, R1’, G1’, B1’, R2’, G2’, B2’, R3’, G3’, B3’
	12	9	2	R0’, G0’, B0’, R1’, G1’, B1’
	16, 16f	6	1	R’, G’, B’
`XYZ`	12	9	2	X0’, Y0’, Z0’, X1’, Y1’, Z1’
	16, 16f	6	1	X’, Y’, Z’

6.2.4 pgroup construction – 4:2:2 Sampling System

When using the 4:2:2 sampling system, pgroups shall be constructed as described in Table 2.

Table 2 — Construction of 4:2:2 pgroups

sampling	depth	pgroup size (octets)	pgroup coverage (pixels)	Sample Order
`YCbCr-4:2:2` **`CLYCbCr-4:2:2`**	8	4	2	C'_B,Y0',C'_R,Y1'
	10	5	2	C'_B,Y0',C'_R,Y1'
	12	6	2	C'_B,Y0',C'_R,Y1'
	16, 16f	8	2	C'_B,Y0',C'_R,Y1'
`ICtCp-4:2:2`	8	4	2	C'_T,I0',C'_P,I1'
	10	5	2	C'_T,I0',C'_P,I1'
	12	6	2	C'_T,I0',C'_P,I1'
	16, 16f	8	2	C'_T,I0',C'_P,I1'

Figure 3 shows an example of the 4:2:2 10-bit case.

Note: The subsampling details and co-siting of samples are defined in the applicable signal definition corresponding to the system colorimetry as indicated by the colorimetry parameter in the SDP.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    C’B (10 bits)  |   Y0’ (10 bits)  |    C’R (10 bits)  |    Y1’ (10 bits)   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3 — Example of a 5 Octet 4:2:2 10-bit pgroup

6.2.5 pgroup construction – 4:2:0 Sampling System

The 4:2:0 sampling system shall only be applied to progressive scan images transmitted in a progressive manner. This sampling system does not apply to PsF or interlaced video essence.

When using the 4:2:0 sampling system, pgroups shall be constructed as described in Table 3 using the numbering illustrated in Figure 4.

Note 1: In interpreting Figure 4, the subscripted indices of the Y and C components reflect the position of the sample within its own sample array, noting that the Y and C sample arrays are of different dimensions due to sub-sampling – so for example the C_B'12 sample is correspondingly located with the Y'24 sample.

Note 2: The subsampling details and co-siting of samples are defined in the applicable signal definition corresponding to the system colorimetry as indicated by the colorimetry parameter in the SDP.

Table 3 — Construction of 4:2:0 pgroups

sampling	depth	pgroup size (octets)	pgroup coverage (pixels)	Sample Order
`YCbCr-4:2:0` `CLYCbCr-4:2:0`	8	6	4	Y'00-Y'01-Y'10-Y'11-C_B'00-C_R'00
	10	15	8	Y'00-Y'01-Y'10-Y'11-C_B'00-C_R'00, Y'02-Y'03-Y'12-Y'13-C_B'01-C_R'01
	12	9	4	Y'00-Y'01-Y'10-Y'11-C_B'00-C_R'00
`ICtCp-4:2:0`	8	6	4	I00-I01-I10-I11-C_T00-C_P00
	10	15	8	I00-I01-I10-I11-C_T00-C_P00, I02-I03-I12-I13-C_T01-C_P01
	12	9	4	I00-I01-I10-I11-C_T00-C_P00

For $Y'C'_BC'_R$ and IC_TC_P 4:2:0 video, the color difference components are sub-sampled by a factor of two both horizontally and vertically. Therefore, color difference samples are shared between certain adjacent luminance sample rows. Figure 4 illustrates the sample numbering relationship between the luminance and color difference samples for $Y'C'_BC'_R$ – the numbering for IC_TC_P follows the same principle.

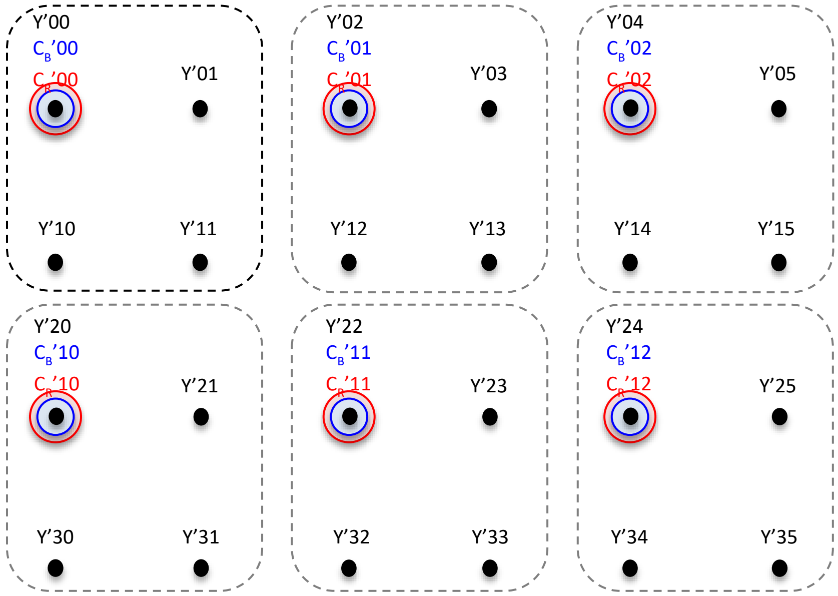

Figure 4 — Sample Numbering in the 4:2:0 system

When packetizing progressive scan 4:2:0 $Y'C'_BC'_R$ or IC_TC_P video data segments, samples from two consecutive luminance sample rows shall be included in each pgroup. The sample row number in the Sample Row Data Header shall be set to that of the first luminance sample row of the pair:

Figure 5 shows an example of the 15 octet pgroup for 4:2:0 10-bit $Y'C'_BC'_R$ format.

```
 0                   1                   2                   3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |   Y’00 (10 bits)   |   Y’01 (10 bits)   |   Y’10 (10 bits)   |   Y’11 (10 bits)   |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |  CB’00 (10 bits)   |  CR’00 (10 bits)   |   Y’02 (10 bits)   |   Y’03 (10 bits)   |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |   Y’12 (10 bits)   |   Y’13 (10 bits)   |  CB’01 (10 bits)   |  CR’01 (10 bits)   |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5 — Example of a 4:2:0 10-bit pgroup

6.2.6 pgroup construction – Key Signal

Key Signals (also known as Alpha Signals) are represented using a single component, "Key". The Key signal is represented using the signaled `depth`, with the signaled `width`, `height`, and `exactframerate` values. The Key signal shall be as defined in SMPTE RP 157. The pgroups of the Key signal shall be constructed as described in Table 4.

Table 4 — Construction of Key signal pgroups

sampling	depth	pgroup size (octets)	pgroup coverage (pixels)	Sample Order
Key	8	1	1	K
	10	5	4	K0, K1, K2, K3
	12	3	2	K0, K1
	16, 16f	2	1	K

6.3 Additional Constraints on the RTP Payload Definition

6.3.1 Packing Modes

The provisions above allow a wide variety of potential practices for mapping active video signals into RTP datagrams. In order to promote interoperability, further constraints on the mapping of active video samples into RTP are documented through the specification of packing modes in sections 6.3.2 and 6.3.3.

Senders conforming to this standard shall operate in one of these packing modes, and shall signal their mode through the use of the appropriate Media Type parameter as indicated. Receivers compliant to this standard shall be capable of receiving all of the listed packing modes.

6.3.2 General Packing Mode (GPM)

This packing mode is suitable for general applications of this standard.

When operating in this mode, the Line Continuation ("C") bit may be used in order to pack samples from more than one sample row into the current packet, to avoid making packets that are too small. IP Datagrams of less than 1000 octets should be avoided except at the end of fields or frames. The sender should make RTP packets which are close to the prevailing UDP Size Limit (Standard or Extended) defined in SMPTE ST 2110-10. The last packet of a field or frame may be padded with one or more octets of value zero.

Senders compliant to this packing mode shall signal compliance with a `PM` Media Type Parameter with value `2110GPM`.

6.3.3 Block Packing Mode (BPM)

This mode is a constrained subset of the General Packing Mode (GPM) defined in section 6.3.2, such that the sum of the lengths of the Sample Row Data Segments shall be a multiple of 180 octets. The largest multiple of 180 consistent with the prevailing maximum UDP size limit shall be used. When operating in this mode, the Line Continuation (‘C’) bit shall be used in order to pack samples from more than one sample row into the current packet, in order to maintain a consistent number of 180 octet blocks per packet. In accordance with ST 2110-10 Section 5, and in consideration of the 12-octet (minimum) RTP header, the maximum available payload space is 1428 octets. Accordingly, a payload of 7 x 180 octets per packet shall be used. The last packet of a field or frame shall not be subject to the 180 octet multiple rule above, and may be either truncated or padded with zeros resulting in the same sized packet as those preceding it. Informative Annex A contains a table of the block sizes consistent with the 7x180 octet payload size above.

The Extended UDP size limit defined in SMPTE ST 2110-10 shall not be used in the Block Packing Mode.

Senders compliant to this mode shall signal compliance with a `PM` Media Type Parameter with value `2110BPM`.

7 Session Description Protocol (SDP) Considerations

7.1 General SDP Declaration

Senders and Receivers shall adhere to IETF RFC 4566 which describes how to map the Media Type Parameters below into the syntax of the SDP object.

The video streams described in this standard shall be declared in the SDP using the Media Type name `video` and the Media Subtype name `raw`. The `rtpmap` clause of the SDP shall indicate the 90 kHz RTP Clock rate.

The SDP metadata items below shall apply equally to all samples, rows, fields, and frames of a stream. For example, the `depth` parameter applies equally to every pgroup in the stream.

IETF RFC 4566 section “SDP Attributes” specifies the `a=fmtp` clause including `<format specific parameters>`. The `<format specific parameters>` section shall consist of a sequence of media type parameter entries, separated by the semicolon (“;”) character followed by whitespace. There is no semicolon character after the last item. The `a=fmtp` clause shall be terminated by a carriage return.

Each media type parameter entry shall be constructed as either

- a `<name>=<value>` pair, with no whitespace within the name or value or between the name, equal sign, and value
- a `<name>` standalone declaration, with no whitespace within the name.

7.2 Required Media Type Parameters

Senders shall include the following payload-format-specific Media Type parameters in the `a=fmtp` clause of the SDP for all streams conforming to this standard.

`sampling`	Signals the color difference signal sub-sampling structure. See section 7.4.1 for permitted values of this parameter under this standard.
`depth`	Signals the number of bits per sample. See section 7.4.2 for permitted values of this parameter.
`width`	Signals the number of pixels per row. Permitted values are integers between 1 and 32767 inclusive.
`height`	Signals the number of full-bandwidth Sample Rows per frame. Permitted values are integers between 1 and 32767 inclusive.
`exactframerate`	Signals the frame rate in frames per second. Integer frame rates shall be signaled as a single decimal number (e.g. “25”) whilst non-integer frame rates shall be signaled as a ratio of two integer decimal numbers separated by a “forward-slash” character (e.g. “30000/1001”), utilizing the numerically smallest numerator value possible.
`colorimetry`	Specifies the system colorimetry used by the image samples. See section 7.5 for permitted values.
`PM`	Packing Mode. Values of this parameter are defined in section 6.3.
`SSN`	SMPTE Standard Number. Senders implementing this standard shall signal the value `ST2110-20:2017` unless the `colorimetry` value `ALPHA` or the `TCS` value `ST2115LOGS3` are used, in which case the value `ST2110-20:2022` shall be signaled.

7.3 Media Type Parameters with default values

Senders shall include the following payload-format-specific Media Type parameters in the `a=fmtp` clause of the SDP for any streams conforming to this standard where the default values are not correct for the contents of the stream.

`interlace`	If this parameter name is present, it indicates that the video is interlaced, or that the video is Progressive segmented Frame (PsF). If this parameter name is not present, the progressive video format shall be assumed.
`segmented`	If this parameter name is present, and the `interlace` parameter name is also present, then the video is a Progressive segmented Frame (PsF). Signaling of this parameter without the `interlace` parameter is forbidden.
`TCS`	Transfer Characteristic System. This parameter specifies the transfer characteristic system of the image samples. See section 7.6 for permitted values of this parameter, and the default assumption of receivers if it is not signaled.

`RANGE` This parameter should be used to signal the encoding range of the sample values within the stream. When paired with Recommendation ITU-R BT.2100 colorimetry, this parameter has two allowed values `NARROW` and `FULL`, corresponding to the ranges specified in table 9 of Recommendation ITU-R BT.2100. In any other context, this parameter has three allowed values: `NARROW`, `FULLPROTECT`, and `FULL`, which correspond to the ranges specified in SMPTE RP 2077 (where FULLPROTECT indicates the "permitted" range specified in RP 2077:2013 section 5). In the absence of this parameter, `NARROW` shall be the assumed value in either case.

`MAXUDP` This parameter specifies the Maximum UDP Packet Size, as defined in SMPTE ST 2110-10. If absent, it indicates that the Standard UDP Size Limit is in use.

`PAR` This parameter specifies the Pixel Aspect Ratio. When it is signaled, `PAR` shall be signaled as a ratio of two integer decimal numbers separated by a "colon" character (e.g. "12:11"). The first integer in the `PAR` is the width of a luminance sample, and the second integer is the height. The smallest integer values possible for width and height shall be used. If `PAR` is not signaled, the receiver shall assume that PAR = "1:1".

7.4 Sampling and Depth

7.4.1 Sampling

Video streams shall use one of the values for the Media Type Parameter `sampling` defined in this section.

Signals utilizing the non-constant luminance $Y'C'_BC'_R$ signal format of Recommendation ITU-R BT.601, Recommendation ITU-R BT.709, Recommendation ITU-R BT.2020, or Recommendation ITU-R BT.2100 shall use the appropriate one of the following values for the Media Type Parameter `sampling`:

`YCbCr-4:4:4` (4:4:4 sampling)

`YCbCr-4:2:2` (4:2:2 sampling)

`YCbCr-4:2:0` (4:2:0 sampling)

Signals utilizing the Constant Luminance $Y'_CC'_{BC}C'_{RC}$ signal format of Recommendation ITU-R BT.2020 shall use the appropriate one of the following values for the Media Type Parameter `sampling`:

`CLYCbCr-4:4:4` (4:4:4 sampling)

`CLYCbCr-4:2:2` (4:2:2 sampling)

`CLYCbCr-4:2:0` (4:2:0 sampling)

Signals utilizing the constant intensity IC_TC_P signal format of Recommendation ITU-R BT.2100 shall use the appropriate one of the following values for the Media Type Parameter `sampling`:

`ICtCp-4:4:4` (4:4:4 sampling)

`ICtCp-4:2:2` (4:2:2 sampling)

`ICtCp-4:2:0` (4:2:0 sampling)

Signals utilizing the 4:4:4 R'G'B' or RGB signal format (such as that of Recommendation ITU-R BT.601, Recommendation ITU-R BT.709, Recommendation ITU-R BT.2020, Recommendation ITU-R BT.2100, SMPTE ST 2065-1 or ST 2065-3) shall use the following value for the Media Type Parameter `sampling`.

`RGB`	RGB or R'G'B' samples

Signals utilizing the 4:4:4 *X'Y'Z'* signal format (such as defined in SMPTE ST 428-1) shall use the following value for the Media Type Parameter `sampling`.

`XYZ`	*X'Y'Z'* samples

Key signals as defined in SMPTE RP 157 shall use the following value for the Media Type Parameter `sampling`. The Key signal is represented as a single component.

`KEY`	samples of the key signal

Key signals are used in relationship to "fill" signals of video content. The Key signal does not have a specific TCS or Colorimetry value itself; the Key stream shall signal the colorimetry value "ALPHA", and shall not signal a TCS value.

Note: Although the IC_TC_P system uses non-linear signals, the nomenclature does NOT employ the prime symbols, per the guidance of Recommendation ITU-R BT.2100.

7.4.2 Depth

Video streams shall signal the appropriate one of the following values of the `depth` Media Type Parameter:

`8`	8 bit integer samples
`10`	10 bit integer samples
`12`	12 bit integer samples
`16`	16 bit integer samples (such as used in SMPTE ST 2065-3 ADX encoding)
`16f`	16 bit floating point samples (as defined in SMPTE ST 2065-1 and Recommendation ITU-R BT.2100)

7.5 Permitted values of Colorimetry

Video streams shall indicate the colorimetry of the signal using one of the following values for the Media Type Parameter `colorimetry`:

`BT601`	as specified in Recommendation ITU-R BT.601-7
`BT709`	as specified in Recommendation ITU-R BT.709-6
`BT2020`	as specified in Recommendation ITU-R BT.2020-2
`BT2100`	as specified in Recommendation ITU-R BT.2100 Table 2 titled "System colorimetry"

ST2065-1	as specified in SMPTE ST 2065-1 Academy Color Encoding Specification (ACES)
ST2065-3	as specified for Academy Density Exchange Encoding (ADX) in SMPTE ST 2065-3
UNSPECIFIED	colorimetry is not specified and must be manually coordinated between sender and receiver.
XYZ	the CIE 1931 standard colorimetric system as specified in ISO 11664-1
ALPHA	colorimetry value signaled for key signals as specified in SMPTE RP 157

Signals utilizing the Recommendation ITU-R BT.2100 colorimetry should also signal the representational range using the optional parameter `RANGE` defined in section 7.3.

7.6 Permitted Values of TCS (Transfer Characteristic System)

Devices indicating a Transfer Characteristic System shall signal it using the `TCS` Media Type Parameter with one of the following values. If the `TCS` value is not specified, receivers shall assume the value `SDR`, unless the `sampling` keyword indicates the signal is a `KEY` signal, in which case the `TCS` value is not meaningful.

SDR	(Standard Dynamic Range) Video streams of standard dynamic range, that utilize the OETF of Recommendation ITU-R BT.709 or Recommendation ITU-R BT.2020. Such streams shall be assumed to target the EOTF specified in Recommendation ITU-R BT.1886.
PQ	Video streams of high dynamic range video that utilize the Perceptual Quantization system of Recommendation ITU-R BT.2100
HLG	Video streams of high dynamic range video that utilize the Hybrid Log-Gamma system of Recommendation ITU-R BT.2100
LINEAR	Video streams of linear encoded floating-point samples (`depth=16f`), such that all values fall within the range [0..1.0].
BT2100LINPQ	Video Stream of linear encoded floating point samples (`depth=16f`) normalized from PQ as specified in Table 10 “Floating Point Signal Representation" of Recommendation ITU-R BT.2100-0
BT2100LINHLG	Video Stream of linear encoded floating point samples (`depth=16f`) normalized from HLG as specified in Table 10 “Floating Point Signal Representation” of Recommendation ITU-R BT.2100-0
ST2065-1	Video stream of linear encoded floating-point samples (`depth=16f`) as specified in SMPTE ST 2065-1
ST428-1	Video stream utilizing the transfer characteristic specified in SMPTE ST 428-1 section 4.3.

`DENSITY`	Video streams of density encoded samples, such as those defined in SMPTE ST 2065-3.
`ST2115LOGS3`	Video streams of high dynamic range video that utilize the “Camera Log S3” transfer characteristic specified in SMPTE ST 2115.
`UNSPECIFIED`	Video streams whose transfer characteristics are not specified. The transfer characteristics must be manually coordinated between sender and receiver.

7.7 Examples of SDP Media Section (Informative)

The example below indicates that the RTP stream is being sent to UDP destination port 30000, with an RTP dynamic payload type of 112 and an RTP Clock rate of 90000 Hz. The image dimensions are 1280x720, and 10-bit 4:2:2 sampling is used. The frame rate is 60/1.001Hz, the transfer characteristic and colorimetry follow Recommendation ITU-R BT.709 and the General Packing Mode is used.

```
m=video 30000 RTP/AVP 112

a=rtpmap:112 raw/90000

a=fmtp:112 sampling=YCbCr-4:2:2; width=1280; height=720;
exactframerate=60000/1001; depth=10; TCS=SDR; colorimetry=BT709;
PM=2110GPM; SSN=ST2110-20:2017
```

Annex A Table of Typical Values for Block Packing Mode (Informative)

Table A.1 lists packet size values for Block Packing Mode which support the required 1260 octet per packet size.

Table A.1— Block Packing Model packet sizes

Sampling	Bit Depth	Per each pgroup		pgroups per 180 block	180 Blocks per Packet	Pixels per Packet	Octets per Packet
		Octets	Pixels				
4:2:2	8	4	2	45	7	630	1260
4:2:2	10	5	2	36	7	504	1260
4:2:2	12	6	2	30	7	420	1260
4:4:4	8	3	1	60	7	420	1260
4:4:4	10	15	4	12	7	336	1260
4:4:4	12	9	2	20	7	280	1260
4:4:4	16	6	1	30	7	210	1260
4:2:0	8	6	4	30	7	840	1260
4:2:0	10	15	8	12	7	672	1260
4:2:0	12	9	4	20	7	560	1260

Bibliography (Informative)

AES67:2018, AES standard for audio applications of networks - High-performance streaming audio-over-IP interoperability

IETF RFC 4175 Gharai, L. and C. Perkins, "RTP Payload Format for Uncompressed Video", RFC 4175, DOI 10.17487/RFC4175, September 2005, https://www.rfc-editor.org/info/rfc4175

SMPTE ST 2022-6:2012 Transport of High Bit Rate Media Signals over IP Networks (HBRMT)

Video Services Forum (VSF) TR-03 "Transport of Uncompressed Elementary Stream Media over IP", November 2015, http://www.videoservicesforum.org/download/technical_recommendations/VSF_TR-03_2015-11-12.pdf

Video Services Forum (VSF) TR-04 "Utilization of ST 2022-6 Media Flows within a VSF TR-03 Environment", November 2015, http://www.videoservicesforum.org/download/technical_recommendations/VSF_TR-04_2015-11-12.pdf

