



IEEE Standard for a System of Advanced Audio and Video Coding

IEEE Computer Society

Developed by the
Data Compression Standards Committee

IEEE Std 1857.3™-2023
(Revision of IEEE Std 1857.3-2013)

STANDARDS

IEEE Standard for a System of Advanced Audio and Video Coding

Developed by the

Data Compression Standards Committee
of the
IEEE Computer Society

Approved 8 November 2023

IEEE SA Standards Board

Abstract: Storage file formats and RTP payload formats for IEEE Std 1857™ video; IEEE Std 1857.2™ audio; IEEE Std 1857.4™ second generation IEEE Std 1857 video; IEEE Std 1857.5™ mobile speech and audio; IEEE Std 1857.8 second generation audio coding; and IEEE Std 1857.10 third generation video coding are defined in this standard. The storage of video and audio not only uses the existing capabilities of the ISO base media file format, but also defines extensions to support specific features of the IEEE Std 1857, IEEE Std 1857.2, IEEE Std 1857.4, IEEE Std 1857.5, IEEE Std 1857.8, and IEEE Std 1857.10 video and audio codecs. The target applications and services include, but are not limited to, Internet media streaming, IPTV, video conference, video telephony, and video-on-demand.

Keywords: audio, box, file format, IEEE 1857.3™, NAL unit, packetization, RTP payload format, storage, system, track, transmission, video

The Institute of Electrical and Electronics Engineers, Inc.
3 Park Avenue, New York, NY 10016-5997, USA

Copyright © 2024 by The Institute of Electrical and Electronics Engineers, Inc.
All rights reserved. Published 5 April 2024. Printed in the United States of America.

IEEE is a registered trademark in the U.S. Patent & Trademark Office, owned by The Institute of Electrical and Electronics Engineers, Incorporated.

PDF: ISBN 979-8-8557-0614-7 STD26854
Print: ISBN 979-8-8557-0615-4 STDPD26854

IEEE prohibits discrimination, harassment, and bullying.

For more information, visit <https://www.ieee.org/about/corporate/governance/p9-26.html>.

No part of this publication may be reproduced in any form, in an electronic retrieval system or otherwise, without the prior written permission of the publisher.

Important Notices and Disclaimers Concerning IEEE Standards Documents

IEEE Standards documents are made available for use subject to important notices and legal disclaimers. These notices and disclaimers, or a reference to this page (<https://standards.ieee.org/ipr/disclaimers.html>), appear in all standards and may be found under the heading “Important Notices and Disclaimers Concerning IEEE Standards Documents.”

Notice and Disclaimer of Liability Concerning the Use of IEEE Standards Documents

IEEE Standards documents are developed within IEEE Societies and subcommittees of IEEE Standards Association (IEEE SA) Board of Governors. IEEE develops its standards through an accredited consensus development process, which brings together volunteers representing varied viewpoints and interests to achieve the final product. IEEE standards are documents developed by volunteers with scientific, academic, and industry-based expertise in technical working groups. Volunteers are not necessarily members of IEEE or IEEE SA and participate without compensation from IEEE. While IEEE administers the process and establishes rules to promote fairness in the consensus development process, IEEE does not independently evaluate, test, or verify the accuracy of any of the information or the soundness of any judgments contained in its standards.

IEEE makes no warranties or representations concerning its standards, and expressly disclaims all warranties, express or implied, concerning this standard, including but not limited to the warranties of merchantability, fitness for a particular purpose and non-infringement. IEEE Standards documents do not guarantee safety, security, health, or environmental protection, or guarantee against interference with or from other devices or networks. In addition, IEEE does not warrant or represent that the use of the material contained in its standards is free from patent infringement. IEEE Standards documents are supplied “AS IS” and “WITH ALL FAULTS.”

Use of an IEEE standard is wholly voluntary. The existence of an IEEE standard does not imply that there are no other ways to produce, test, measure, purchase, market, or provide other goods and services related to the scope of the IEEE standard. Furthermore, the viewpoint expressed at the time a standard is approved and issued is subject to change brought about through developments in the state of the art and comments received from users of the standard.

In publishing and making its standards available, IEEE is not suggesting or rendering professional or other services for, or on behalf of, any person or entity, nor is IEEE undertaking to perform any duty owed by any other person or entity to another. Any person utilizing any IEEE Standards document, should rely upon their own independent judgment in the exercise of reasonable care in any given circumstances or, as appropriate, seek the advice of a competent professional in determining the appropriateness of a given IEEE standard.

IN NO EVENT SHALL IEEE BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO: THE NEED TO PROCURE SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE PUBLICATION, USE OF, OR RELIANCE UPON ANY STANDARD, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE AND REGARDLESS OF WHETHER SUCH DAMAGE WAS FORESEEABLE.

Translations

The IEEE consensus balloting process involves the review of documents in English only. In the event that an IEEE standard is translated, only the English version published by IEEE is the approved IEEE standard.

Official statements

A statement, written or oral, that is not processed in accordance with the IEEE SA Standards Board Operations Manual shall not be considered or inferred to be the official position of IEEE or any of its committees and shall not be considered to be, nor be relied upon as, a formal position of IEEE. At lectures, symposia, seminars, or educational courses, an individual presenting information on IEEE standards shall make it clear that the presenter's views should be considered the personal views of that individual rather than the formal position of IEEE, IEEE SA, the Standards Committee, or the Working Group. Statements made by volunteers may not represent the formal position of their employer(s) or affiliation(s).

Comments on standards

Comments for revision of IEEE Standards documents are welcome from any interested party, regardless of membership affiliation with IEEE or IEEE SA. However, **IEEE does not provide interpretations, consulting information, or advice pertaining to IEEE Standards documents.**

Suggestions for changes in documents should be in the form of a proposed change of text, together with appropriate supporting comments. Since IEEE standards represent a consensus of concerned interests, it is important that any responses to comments and questions also receive the concurrence of a balance of interests. For this reason, IEEE and the members of its Societies and subcommittees of the IEEE SA Board of Governors are not able to provide an instant response to comments, or questions except in those cases where the matter has previously been addressed. For the same reason, IEEE does not respond to interpretation requests. Any person who would like to participate in evaluating comments or in revisions to an IEEE standard is welcome to join the relevant IEEE working group. You can indicate interest in a working group using the Interests tab in the Manage Profile and Interests area of the [IEEE SA myProject system](#).¹ An IEEE Account is needed to access the application.

Comments on standards should be submitted using the [Contact Us](#) form.²

Laws and regulations

Users of IEEE Standards documents should consult all applicable laws and regulations. Compliance with the provisions of any IEEE Standards document does not constitute compliance to any applicable regulatory requirements. Implementers of the standard are responsible for observing or referring to the applicable regulatory requirements. IEEE does not, by the publication of its standards, intend to urge action that is not in compliance with applicable laws, and these documents may not be construed as doing so.

Data privacy

Users of IEEE Standards documents should evaluate the standards for considerations of data privacy and data ownership in the context of assessing and using the standards in compliance with applicable laws and regulations.

¹Available at: <https://development.standards.ieee.org/myproject-web/public/view.html#landing>.

²Available at: <https://standards.ieee.org/content/ieee-standards/en/about/contact/index.html>.

Copyrights

IEEE draft and approved standards are copyrighted by IEEE under US and international copyright laws. They are made available by IEEE and are adopted for a wide variety of both public and private uses. These include both use, by reference, in laws and regulations, and use in private self-regulation, standardization, and the promotion of engineering practices and methods. By making these documents available for use and adoption by public authorities and private users, neither IEEE nor its licensors waive any rights in copyright to the documents.

Photocopies

Subject to payment of the appropriate licensing fees, IEEE will grant users a limited, non-exclusive license to photocopy portions of any individual standard for company or organizational internal use or individual, non-commercial use only. To arrange for payment of licensing fees, please contact Copyright Clearance Center, Customer Service, 222 Rosewood Drive, Danvers, MA 01923 USA; +1 978 750 8400; <https://www.copyright.com/>. Permission to photocopy portions of any individual standard for educational classroom use can also be obtained through the Copyright Clearance Center.

Updating of IEEE Standards documents

Users of IEEE Standards documents should be aware that these documents may be superseded at any time by the issuance of new editions or may be amended from time to time through the issuance of amendments, corrigenda, or errata. An official IEEE document at any point in time consists of the current edition of the document together with any amendments, corrigenda, or errata then in effect.

Every IEEE standard is subjected to review at least every 10 years. When a document is more than 10 years old and has not undergone a revision process, it is reasonable to conclude that its contents, although still of some value, do not wholly reflect the present state of the art. Users are cautioned to check to determine that they have the latest edition of any IEEE standard.

In order to determine whether a given document is the current edition and whether it has been amended through the issuance of amendments, corrigenda, or errata, visit [IEEE Xplore](#) or [contact IEEE](#).³ For more information about the IEEE SA or IEEE's standards development process, visit the IEEE SA Website.

Errata

Errata, if any, for all IEEE standards can be accessed on the [IEEE SA Website](#).⁴ Search for standard number and year of approval to access the web page of the published standard. Errata links are located under the Additional Resources Details section. Errata are also available in [IEEE Xplore](#). Users are encouraged to periodically check for errata.

Patents

IEEE standards are developed in compliance with the [IEEE SA Patent Policy](#).⁵

Attention is called to the possibility that implementation of this standard may require use of subject matter covered by patent rights. By publication of this standard, no position is taken by the IEEE with respect to the existence or validity of any patent rights in connection therewith. If a patent holder or patent applicant has

³Available at: <https://ieeexplore.ieee.org/browse/standards/collection/ieee>.

⁴Available at: <https://standards.ieee.org/standard/index.html>.

⁵Available at: <https://standards.ieee.org/about/sasb/patcom/materials.html>.

filed a statement of assurance via an Accepted Letter of Assurance, then the statement is listed on the IEEE SA Website at <https://standards.ieee.org/about/sasb/patcom/patents.html>. Letters of Assurance may indicate whether the Submitter is willing or unwilling to grant licenses under patent rights without compensation or under reasonable rates, with reasonable terms and conditions that are demonstrably free of any unfair discrimination to applicants desiring to obtain such licenses.

Essential Patent Claims may exist for which a Letter of Assurance has not been received. The IEEE is not responsible for identifying Essential Patent Claims for which a license may be required, for conducting inquiries into the legal validity or scope of Patents Claims, or determining whether any licensing terms or conditions provided in connection with submission of a Letter of Assurance, if any, or in any licensing agreements are reasonable or non-discriminatory. Users of this standard are expressly advised that determination of the validity of any patent rights, and the risk of infringement of such rights, is entirely their own responsibility. Further information may be obtained from the IEEE Standards Association.

IMPORTANT NOTICE

Technologies, application of technologies, and recommended procedures in various industries evolve over time. The IEEE standards development process allows participants to review developments in industries, technologies, and practices, and to determine what, if any, updates should be made to the IEEE standard. During this evolution, the technologies and recommendations in IEEE standards may be implemented in ways not foreseen during the standard's development. IEEE standards development activities consider research and information presented to the standards development group in developing any safety recommendations. Other information about safety practices, changes in technology or technology implementation, or impact by peripheral systems also may be pertinent to safety considerations during implementation of the standard. Implementers and users of IEEE Standards documents are responsible for determining and complying with all appropriate safety, security, environmental, health, and interference protection practices and all applicable laws and regulations.

Participants

At the time this standard was completed, the Audio and Video Working Group had the following membership:

Siwei Ma, *Chair*
Cliff Reader, *Vice Chair*

At the time this standard was completed, the 1857.3 Sub-Working Group had the following membership:

Luntian Mou, *Chair*
Haiwu Zhao, *Vice Chair*
Xiaoxu Luan, *Secretary*
Ying Hu, *Technical Editor*

Xilin Chen
Qiang Fu
Cheng Huang

Tiejun Huang
Longshe Huo

Qiuting Li
Jun Sun
Yiling Xu

The following members of the individual Standards Association balloting group voted on this standard. Balloters may have voted for approval, disapproval, or abstention.

Boon Chong Ang
Eric Chai
Pin Chang
Werner Hoelzl
Piotr Karocki
Alexis Kokkonos
Jun Li
Li Li

Weiping Li
Fan Liang
Changyu Liu
Dong Liu
Qian Liu
Xiaoxu Luan
Luntian Mou

Honggang Qi
Yonghong Tian
John Vergis
Yaowei Wang
Feng Wu
Shumin Xu
Oren Yuen
Wenwu Zhu

When the IEEE SA Standards Board approved this standard on 8 November 2023, it had the following membership:

David J. Law, *Chair*
Ted Burse, *Vice Chair*
Gary Hoffman, *Past Chair*
Konstantinos Karachalios, *Secretary*

Sara R. Biyabani
Doug Edward
Ramy Ahmed Fathy
Guido R. Hiertz
Yousef Kimiagar
Joseph L. Koepfinger
Thomas Koshy
John D. Kulick

Joseph S. Levy
Howard Li
Johnny Daozhuang Lin
Gui Lin
Xiaohui Liu
Kevin W. Lu
Daleep C. Mohla
Andrew Myles

Paul Nikolich
Annette D. Reilly
Robby Robson
Lei Wang
F. Keith Waters
Karl Weber
Philip B. Winston
Don Wright

Introduction

This introduction is not part of IEEE Std 1857.3-2023, IEEE Standard for System of Advanced Audio and Video Coding.

This standard defines the storage file formats and RTP payload formats for IEEE Std 1857 video, IEEE Std 1857.2 audio, IEEE Std 1857.4 second generation IEEE Std 1857 video, IEEE Std 1857.5 mobile speech and audio, IEEE Std 1857.8 second generation audio coding, and IEEE Std 1857.10 third generation video coding.

Contents

1. Overview	10
1.1 Scope	10
1.2 Purpose	10
1.3 Word usage	10
2. Normative references	11
3. Definitions, acronyms, and abbreviations	11
3.1 Definitions	11
3.2 Acronyms and abbreviations	12
4. File format for IEEE 1857 video, IEEE 1857.4 video, and IEEE 1857.10 video	13
4.1 NAL unit	13
4.2 IEEE 1857 video, IEEE 1857.4 video, and IEEE 1857.10 video elementary stream and sample definitions	15
4.3 Metadata for surveillance video	26
5. RTP payload format for IEEE 1857 video, IEEE 1857.4 video, and IEEE 1857.10 video	29
5.1 NAL unit	29
5.2 RTP payload format	29
5.3 Packetization rules	42
5.4 Payload format parameters	43
6. File format for IEEE 1857.2 audio, IEEE 1857.5 advanced mobile speech and audio, and IEEE 1857.8 audio	53
6.1 NAL unit	53
6.2 IEEE 1857.2 audio, IEEE 1857.5 advanced mobile speech and audio, and IEEE 1857.8 audio elementary stream and sample definition	54
7. RTP payload format for IEEE 1857.2 audio, IEEE 1857.5 advanced mobile speech and audio, and IEEE 1857.8 audio	58
7.1 NAL unit	58
7.2 RTP payload format	58
7.3 Packetization rules	67
7.4 Payload format parameters	68
Annex A (informative) De-packetization process	74

IEEE Standard for a System of Advanced Audio and Video Coding

1. Overview

1.1 Scope

This standard specifies the storage file format and RTP (Real-time Transport Protocol) payload format for the compressed video data created by IEEE Std 1857TM, IEEE Std 1857.4TM, IEEE Std 1857.10TM, and the compressed audio data created by IEEE Std 1857.2TM, IEEE Std 1857.5TM, and IEEE Std 1857.8TM.⁶ The storage format is aimed to extend ISO base media file format to support specific features of IEEE Std 1857 video, IEEE Std 1857.2 audio, IEEE Std 1857.4 second generation IEEE 1857 video coding, IEEE Std 1857.5 mobile speech and audio, IEEE Std 1857.8 second generation audio coding, and IEEE Std 1857.10 third generation video coding, and the RTP payload format defines the specific payload format to convey IEEE 1857 video, IEEE Std 1857.2 audio, IEEE Std 1857.4 video, IEEE Std 1857.5 mobile speech and audio, IEEE Std 1857.8 second generation audio coding, and IEEE Std 1857.10 third generation video coding by RTP packets.

1.2 Purpose

This standard provides regular and efficient system level tool sets that describe the file storage format and RTP payload format for IEEE 1857 video, IEEE 1857.2 audio, IEEE 1857.4 video, IEEE 1857.5 mobile speech and audio, IEEE 1857.8 audio, and IEEE 1857.10 video while storing and streaming over IP network.

1.3 Word usage

The word *shall* indicates mandatory requirements strictly to be followed in order to conform to the standard and from which no deviation is permitted (*shall* equals *is required to*).^{7,8}

The word *should* indicates 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 (*should* equals *is recommended that*).

The word *may* is used to indicate a course of action permissible within the limits of the standard (*may* equals *is permitted to*).

⁶Information on references can be found in [Clause 2](#).

⁷The use of the word *must* is deprecated and cannot be used when stating mandatory requirements, *must* is used only to describe unavoidable situations.

⁸The use of *will* is deprecated and cannot be used when stating mandatory requirements, *will* is only used in statements of fact.

The word *can* is used for statements of possibility and capability, whether material, physical, or causal (*can* equals *is able to*).

2. Normative references

The following referenced documents are indispensable for the application of this document (i.e., they must be understood and used, so each referenced document is cited in text and its relationship to this document is explained). For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments or corrigenda) applies.

IEEE Std 1857™, IEEE Standard for Advanced Audio and Video Coding.^{9,10}

IEEE Std 1857.2™, IEEE Standard for Advanced Audio Coding.

IEEE Std 1857.4™, IEEE Standard for Second Generation IEEE 1857 Video Coding.

IEEE Std 1857.5™, IEEE Standard for Advanced Mobile Speech and Audio.

IEEE Std 1857.8™, IEEE Standard for Second Generation Audio Coding.

IEEE Std 1857.10™, IEEE Standard for Third Generation Video Coding.

IETF RFC 2327, SDP: Session Description Protocol.¹¹

IETF RFC 3264, An Offer/Answer Model with Session Description Protocol (SDP).

IETF RFC 3550, RTP: A Transport Protocol for Real-Time Applications.

ISO/IEC 14496-12, Information Technology—Coding of Audio-Visual Objects—ISO Base Media File Format.

3. Definitions, acronyms, and abbreviations

3.1 Definitions

For the purposes of this document, the following terms and definitions apply. The *IEEE Standards Dictionary Online* should be consulted for terms not defined in this clause.¹²

access unit: A series of network abstract layer (NAL) units that constitute a frame of coded picture.

NOTE—Besides picture header and slice data, an access unit can also contain other types of coding data. All data within the same access unit should have the same timestamp value for real-time transport protocol (RTP) packetization.¹³

⁹The IEEE standards or products referred to in this clause are trademarks owned by The Institute of Electrical and Electronics Engineers, Incorporated.

¹⁰IEEE publications are available from The Institute of Electrical and Electronics Engineers (<https://standards.ieee.org/>).

¹¹RFC documents are available through the IETF at <https://www.ietf.org/standards/rfcs/>.

¹²*IEEE Standards Dictionary Online* is available at: <http://dictionary.ieee.org>. An IEEE account is required for access to the dictionary, and one can be created at no charge on the dictionary sign-in page.

¹³Notes in text, tables, and figures of a standard are given for information only and do not contain requirements needed to implement this standard.

decoding order number (DON): Network abstract layer (NAL) unit decoding order indicated by a field in the real-time transport protocol (RTP) payload structure or a derived variable.

NOTE—Values of DON are in the range from 0 to 65535, inclusive. After reaching the maximum value, the value of DON wraps around to 0.

media aware network element (MANE): A network element capable of parsing certain aspects of the RTP payload headers or the RTP payload and reacting to the contents.

NOTE—Examples of MANE are intelligent middleboxes and application layer gateways. Compared with normal routers or gateways, MANE should know the contents of the signal (e.g., mapping of media stream payload types). Furthermore, MANE should be a trusted entity when Secure Real-time Transport Protocol (SRTP) is in use. MANE has the advantage of losing packets according to the needs of media coding. For example, if MANE should discard packets due to congestion of some link, it can identify and discard the packets that reduce the impact on user experience.

NALU: A network abstract layer unit composed of one byte header and one or more bytes data.

NOTE—NALU for IEEE 1857 video, IEEE 1857.4 video, and IEEE 1857.10 video is defined in 5.1. NALU for IEEE 1857.2 audio, IEEE 1857.5 mobile speech and audio, and IEEE 1857.8 audio is defined in 7.1.

NALU decoding order: The order of NALUs in a network abstraction layer (NAL) unit stream.

NAL unit stream: A sequence composed of one or more network abstraction layer (NAL) units.

packet: Datagram encapsulated using IP protocol.

transmission order: The order of packets in ascending real-time transport protocol (RTP) sequence number order (in modulo arithmetic).

3.2 Acronyms and abbreviations

AVS	audio and video coding standard
DON	decoding order number
DONB	decoding order number base
DOND	decoding order number difference
FU	fragmentation unit
IDR	instantaneous decoding refresh
MANE	media aware network element
MIME	multipurpose internet mail extension
MTAP	multi-time aggregation packet
MTAP16	MTAP with 16-bit timestamp offset
MTAP24	MTAP with 24-bit timestamp offset
MTU	maximum transfer unit
NAL	network abstraction layer
NALU	NAL unit
PSI	payload structure indicator

RTP	real-time transport protocol
STAP	single-time aggregation packet
STAP-A	STAP type A
STAP-B	STAP type B

4. File format for IEEE 1857 video, IEEE 1857.4 video, and IEEE 1857.10 video

4.1 NAL unit

The conversion process from video bit stream defined in IEEE 1857 video, IEEE 1857.4 video, or IEEE 1857.10 video to NAL unit stream is as follows:

- Map the data between every two consecutive start code prefixes (0x000001) in the video bit stream (including start code values but excluding start code prefixes) into the NAL unit data.
- Insert a NAL unit header before the NAL unit data according to the start code value and its context.

The format of a NAL unit header is shown in [Figure 1](#).

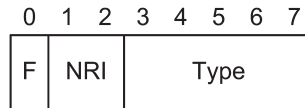


Figure 1—The format of video NALU header

The syntax and semantics of the NAL unit header are as follows:

- **F:** This 1-bit field is for forbidden zero bit. Its value should be 0.
- **NRI:** This 2-bit field is for NAL reference identification (nal_ref_idc). Value of nonzero means that the data contained in this NAL unit is sequence header or reference frame data. Value of 0 means that the data contained in this NAL unit is not reference frame data. For sequence header NAL unit, nal_ref_idc should not be 0. For a certain frame, if nal_ref_idc of one NAL unit is 0, then nal_ref_idc of all NAL units in the same frame should be 0. Nal_ref_idc of NAL units for I frames should not be 0.
- **Type:** This 5-bit field is for NAL unit type (nal_unit_type). The value of this field is decided according to the start code value and the picture header contained in the following NAL unit data, and their context, as shown in [Table 1](#) for IEEE 1857 video, in [Table 2](#) for IEEE 1857.4 video, and in [Table 3](#) for IEEE 1857.10 video.

Table 1—Value assignment for the NAL unit type field in NAL unit header for IEEE 1857 video

Type	Type description	Reason for type assignment
0	Reserved	—
1	Sequence header	Start code value is B0
2	video extension	Start code value is B5
3	User data	Start code value is B2
4	video edit	Start code value is B7
5	Picture header of I frame	Start code value is B3
6	Picture header of P frame	Start code value is B6, and the picture coding type of the picture header is 01
7	Picture header of B frame	Start code value is B6, and the picture coding type of the picture header is 10
8	Slice data of I frame	Start code value is 00~AF, and the start code value of the last picture header before this NALU is B3
9	Slice data of P frame	Start code value is 00~AF, and the start code value of the last picture header before this NALU is B6, and the picture coding type of the picture is 01
10	Slice data of B frame	Start code value is 00~AF, and the start code value of the last picture header before this NALU is B6, and the picture coding type of the picture is 10
11–23	Reserved	—
24–31	Unspecified	—

Table 2—Value assignment for the NAL unit type field in NAL unit header for IEEE 1857.4 video

Type	Type description	Reason for type assignment
0	slice_start_code	Start code value is 00~8F
1	reserved	Start code value is 90~AF
2	video_sequence_start_code	Start code value is B0
3	video_sequence_end_code	Start code value is B1
4	user_data_start_code	Start code value is B2
5	intra_picture_start_code	Start code value is B3
6	reserved	Start code value is B4
7	extension_start_code	Start code value is B5
8	inter_picture_start_code	Start code value is B6
9	video_edit_code	Start code value is B7
10	reserved	Start code value is B8
11	System start code	Start code value is B9~FF
12–23	Reserved	—
24–31	Unspecified	—

Table 3—Value assignment for the NAL unit type field in NAL unit header for IEEE 1857.10 video

Type	Type description	Reason for type assignment
0	patch_start_code	Start code value is 00~7F
1	reserved	Start code value is 80~8E
2	patch_end_code	Start code value is 8F
3	reserved	Start code value is 90~9F
4	video_sequence_start_code	Start code value is B0
5	video_sequence_end_code	Start code value is B1
6	user_data_start_code	Start code value is B2
7	intra_picture_start_code	Start code value is B3
8	reserved	Start code value is B4
9	extension_start_code	Start code value is B5
10	inter_picture_start_code	Start code value is B6
11	video_edit_code	Start code value is B7
12	reserved	Start code value is B8
13	System start code	Start code value is B9~FF
14–23	Reserved	—
24–31	Unspecified	—

Before using the file format defined by this subclause, IEEE 1857 video bit stream, IEEE 1857.4 video bit stream, and IEEE 1857.10 video bit stream should be converted into NAL unit stream at first.

When the decoder receives NAL unit stream, before decoding, it shall discard every NAL unit header of a NAL unit, and then insert a start code prefix (0x000001) in the same position to transform the NAL unit stream back into an IEEE 1857 video bit stream, an IEEE 1857.4 video bit stream, or an IEEE 1857.10 video bit stream.

4.2 IEEE 1857 video, IEEE 1857.4 video, and IEEE 1857.10 video elementary stream and sample definitions

This subclause specifies the elementary streams and sample structures used to store IEEE 1857 video, IEEE 1857.4 video, or IEEE 1857.10 video content inside the file format.

4.2.1 Elementary stream structure

NAL unit stream should be converted from IEEE 1857 video, IEEE 1857.4 video, or IEEE 1857.10 video bit stream as specified in 4.1.

Using video elementary stream, the visual content shall be stored with the following constraints:

- Sequence header NAL units shall be stored in the sample description of this track.
- All NAL units, including sequence header NAL units, shall be present in the samples of this video elementary stream.

Two types of elementary streams are defined for storing IEEE 1857.10 video bitstream:

- a) Sequence elementary stream can be decoded with optional reference to library pictures provided by the information other than in such a bitstream.

- b) Library elementary stream contains only library pictures, which can be referenced by P picture or B picture in other elementary streams and should not be used for display.

4.2.2 Sample and configuration definition

4.2.2.1 Sample definition

IEEE 1857 video sample: An IEEE 1857 video sample contains an IEEE 1857 video access unit.

IEEE 1857.4 video sample: An IEEE 1857.4 video sample contains an IEEE 1857.4 video access unit.

IEEE 1857.10 video sample: An IEEE 1857.10 video sample contains an IEEE 1857.10 video access unit.

4.2.2.2 Canonical order and restrictions

The elementary stream is stored in the ISO base media file format in the following canonical formats so that systems that need to customize the stream for delivery over different transport protocols—MPEG-2 Systems, RTP, and so on—should not have to remove information from the stream while being free to add to the stream:

- **Video data NAL units (coded picture header, coded slice):** The picture header and all slice NAL units for a coded picture shall be contained within the sample whose decoding time and composition time are those of the picture. Each IEEE 1857 video sample, IEEE 1857.4 video sample or IEEE 1857.10 video sample shall contain at least one picture header NAL unit and one slice NAL unit.
- **NAL units sequence:** The sequence of NAL units in an elementary stream and within a single sample shall be in a valid decoding order for those NAL units as specified in IEEE Std 1857, IEEE Std 1857.4, or IEEE Std 1857.10 video sample.
- **Timing information:** All timing information is provided by the information stored in the various sample metadata tables. Timing provided within the IEEE 1857 video stream, IEEE 1857.4 video sample, or IEEE 1857.10 video stream in this file format should be ignored.
- **No start code prefixes:** The elementary streams shall not include start code prefixes. Systems that wish to deliver, from this file format, a stream using start code prefixes will need to reformat the stream back into IEEE 1857 video bit stream, IEEE 1857.4 video bit stream, or IEEE 1857.10 video stream by inserting those start code prefixes and discarding those NAL unit headers.

4.2.2.3 Sample structure definition

An IEEE Std 1857 video, IEEE Std 1857.4 video, or IEEE Std 1857.10 video access unit is made up of one or a set of NAL units. Each NAL unit is represented with the length of NAL unit and NAL unit itself:

- *Length:* Indicates the length in bytes of the following NAL unit. The length field can be configured to be of 1, 2, or 4 B.
- *NAL unit:* Contains a NAL unit header followed by NAL unit data as specified in 4.1.

For syntax and semantics of the sample structure, please refer to 4.2.3.3.2.

4.2.2.4 Decoder configuration information

Decoder configuration record contains the size of the length field used in each sample as well as the initial sequence header.

This record contains a version field. This version of this specification defines Version 1 of this record. Incompatible changes to the record will be indicated by a change of version number. Readers shall not attempt to decode this record or the streams to which it applies if the version number is unrecognized.

Compatible extensions to this record will not change the configuration version code. Readers should be prepared to ignore unrecognized data beyond the definition of the data they understand (e.g., after the initial sequence header in this part).

The values of **AvsProfileIndication**, **AvsLevelIndication** shall be valid for all sequence header of the stream described by this record. **AvsLevelIndication** shall indicate a level of capability equal to or greater than the highest level indicated in the included sequence header. **AvsProfileIndication** shall indicate a profile to which the entire stream conforms. If the sequence headers are marked with different profiles, then the stream may need examination to determine to which profile, if any, the stream conforms. If the stream is not examined, or the examination reveals that there is no profile to which the stream conforms, then the stream shall be split into two or more sub-streams with separate configuration records in which these rules can be met.

Syntax and semantics description of decoder configuration record are as follows:

```
aligned(8) class AvsDecoderConfigurationRecord
{
    unsigned int(8) configurationVersion = 1;
    unsigned int(8) AvsProfileIndication;
    unsigned int(8) AvsLevelIndication;
    bit(6) reserved = '111111'b;
    unsigned int(2) lengthSizeMinusOne;
    bit(3) reserved = '111'b;
    unsigned int(5) numOfSequenceHeader;
    for (i=0; i< numOfSequenceHeader; i++)
    {
        unsigned int(16) SequenceHeaderLength ;
        bit(8× SequenceHeaderLength) SequenceHeaderNALUnit;
    }
}
```

AvsProfileIndication contains the profile code as defined in the IEEE 1857 video or IEEE 1857.4 video.

AvsLevelIndication contains the level code as defined in the IEEE 1857 video or IEEE 1857.4 video.

lengthSizeMinusOne indicates the length in bytes of the **NALUnitLength** field in an IEEE 1857 video sample or an IEEE 1857.4 video sample minus one. For example, a size of one byte is indicated with a value of 0. The value of this field shall be one of 0, 1, or 3 corresponding to a length encoded with 1, 2, or 4 B, respectively.

numOfSequenceHeader indicates the number of initial sequence header sets for decoding the elementary stream.

SequenceHeaderLength indicates the length in bytes of the sequence header NAL unit.

SequenceHeaderNALUnit contains a sequence header NAL unit.

Avs3DecoderConfigurationRecord shall apply to IEEE 1857.10 video bitstream that does not support temporal scalability or containing the lowest temporal layer. The decoder configuration record is used for IEEE 1857.10 video bitstream, which needs referencing one or more lower temporal layers for decoding.

Syntax and semantics description of decoder configuration record are as follows:

```
aligned(8) class Avs3DecoderConfigurationRecord {
    unsigned int(8) configurationVersion = 1;
    unsigned int(16) sequence_header_length;
    bit(8×sequence_header_length) sequence_header;
    bit(6) reserved = '111111'b;
    unsigned int(2) library_dependency_idc;
}
```

- a) **sequence_header_length** indicates the length in bytes of the sequence_header() of elementary stream to which this configuration record applies.
- b) **sequence_header** contains a sequence_header() as defined in IEEE Std 1857.10 with the following constraints:
 - 1) The value of library_stream_flag field in the sequence_header() shall be set to '0' if this configuration record applies to the sequence elementary stream. The value of library_stream_flag field in the sequence_header() shall be set to '1' if this configuration record applies to the library elementary stream.
 - 2) The value of library_picture_enable_flag field in the sequence_header() shall be set to '0' if this configuration record applies to the sequence elementary stream which does not refer to library pictures in the library elementary stream for decoding. The value of library_picture_enable_flag field in the sequence_header() shall be set to '1' if this configuration record applies to the sequence elementary stream that refers to the library pictures in the library elementary stream for decoding.
- c) **library_dependency_idc** indicates the type of the elementary stream to which this configuration record applies. If applicable to the sequence elementary stream, library_dependency_idc indicates whether the sequence elementary stream shall be decoded with reference to library pictures in the library elementary stream. The value and description of library_dependency_idc are shown in Table 4.

Table 4—Value and description of library_dependency_idc

Value	Description
00	The configuration record applies to the sequence elementary stream which could be decoded without reference to the library pictures in the library elementary stream.
01	The configuration record applies to the sequence elementary stream which shall be decoded with reference to the library pictures in the library elementary stream.
10	The configuration record applies to the library elementary stream.
11	Both the sequence elementary stream and the library elementary stream are encapsulated in the track corresponding to the configuration record, and the sequence elementary stream shall be decoded with reference to the library pictures in the library elementary stream.

LAvs3DecoderConfigurationRecord shall apply to video content stored in multiple different IEEE 1857.10 tracks, each of which contains one or more upper temporal layers.

```
aligned(8) class LAvs3DecoderConfigurationRecord {
    unsigned int(8) configurationVersion = 1;
    unsigned int(8) num_temporal_layers;
    for(i=0; i< num_temporal_layers; i++) {
        unsigned int(3) temporal_layer_id[i];
        unsigned int(4) frame_rate_code[i];
        bit(1) reserved = '1'b;
        unsigned int(18) temporal_bit_rate_lower[i];
        unsigned int(12) temporal_bit_rate_upper[i];
        bit(2) reserved = '11'b;
    }
}
```

num_temporal_layers indicates the number of temporal layers contained in the elementary stream to which this configuration record applies.

temporal_layer_id[i] indicates the temporal layer ID contained in the elementary stream to which this configuration record applies.

frame_rate_code[i] indicates the frame rate code corresponding to the IEEE 1857.10 video bitstream when the accumulated temporal layer is equal to temporal_layer_id[i].

temporal_bit_rate_lower[i] indicates the lower 18-bits of the corresponding bit rate of the IEEE 1857.10 video bitstream when the accumulated temporal layer is equal to temporal_layer_id[i].

temporal_bit_rate_upper[i] indicates the upper 12-bits of the corresponding bit rate of the IEEE 1857.10 video sub-bit stream when the accumulated temporal layer is equal to temporal_layer_id[i].

4.2.3 Derivation from ISO base media file format

4.2.3.1 File type and identification

The brand “avs2” (for IEEE 1857 video), “avst” (IEEE 1857.4 video) or “avs3” (IEEE 1857.10 video) indicates a file is compliant with the extension to ISO base media file format as specified in this subclause. The brands “avs2,” “avst,” or “avs3” can be defined as a major brand by certain specification. In this case, file extension name and features required are defined by that specification.

Conformance with this specification is indicated by the presence of the brand of a specification that permits the inclusion of IEEE 1857 video, IEEE 1857.4 video, or IEEE 1857.10 video content, in the compatible brands list of the FileTypeBox as defined in the ISO base media file format specification. The file extension name is defined as “.asm” (AVS streaming media).

Readers conforming to this specification should read the file if a suitable brand occurs in the compatible-brands list. Other structures and/or track types, defined in specifications other than that identified by the brand, may be present, and these may be ignored by a reader conformant with the specification identified by the brand.

4.2.3.2 Track structure

In the terminology of the ISO base media file format specification, tracks that store IEEE 1857 video, IEEE 1857.4 video, or IEEE 1857.10 video are video tracks with the following constraints:

- A handler_type of 'vide' in the HandlerBox
- A video media header 'vmhd'
- A sample entry of VisualSampleEntry

4.2.3.3 IEEE 1857 video stream definition

4.2.3.3.1 Sample entry

4.2.3.3.1.1 Definition

Box Types: 'avs2,' 'avss'

Container: Sample Table Box ('stbl')

Mandatory: The 'avs2' box is mandatory

Quantity: One or more sample entries may be present

An IEEE 1857 video stream sample entry shall contain an AvsSequenceInfoBox, including a decoder configuration record AvsDecoderConfigurationRecord.

An optional BitRateBox may be present in the sample entry to signal the bit rate information of the video stream.

Multiple sample descriptions may be used, as permitted by the ISO base media file format specification, to indicate sections of video that use different decode configurations or sequence header.

4.2.3.3.1.2 Syntax

```
class AvsSequenceInfoBox extends Box('avss')
{
    AvsDecoderConfigurationRecord AvsConfig;
}
class BitRateBox extends Box('btrt')
{
    unsigned int(32) maxBitRate;
    unsigned int(32) avgBitRate;
}
class AvsSampleEntry() extends VisualSampleEntry ('avs2'){
    AvsSequenceInfoBox SeqInfo;
    BitRateBox(); // optional
}
```

4.2.3.3.1.3 Semantics

Compressorname in the base class VisualSampleEntry indicates the name of the compressor used with the value “\016AVS1-P2 Coding” being recommended (\016 is 14, the length of the string as a byte).

maxBitrate gives the maximum rate in bits/second over any window of one second.

avgBitrate gives the average rate in bits/second over the entire presentation.

4.2.3.3.2 Sample format

4.2.3.3.2.1 Syntax

```
class AvsSample
{
    unsigned int PictureLength = sample_size;
    for (i=0; i<PictureLength;)
    {
        unsigned int((AvsDecoderConfigurationRecord.lengthSizeMinusOne+1)*8)
        NALUnitLength;
        bit(NALUnitLength * 8) NALUnit;
        i += (AvsDecoderConfigurationRecord.lengthSizeMinusOne+1) +
        NALUnitLength;
    }
}
```

4.2.3.3.2.2 Semantics

NALUnitLength indicates the size of a NAL unit measured in bytes, not including the length field itself.

NALUnit contains a single NAL unit.

4.2.3.4 IEEE 1857.4 video stream definition

4.2.3.4.1 Sample entry

4.2.3.4.1.1 Definition

Box Types: ‘avst,’ ‘avss’

Container: Sample Table Box (‘stbl’)

Mandatory: The ‘avst’ box is mandatory

Quantity: One or more sample entries may be present

An IEEE 1857.4 video stream sample entry shall contain an AvsSequenceInfoBox, including a decoder configuration record AvsDecoderConfigurationRecord.

An optional BitRateBox may be present in the sample entry to signal the bit rate information of the video stream.

Multiple sample descriptions may be used, as permitted by the ISO base media file format specification, to indicate sections of video that use different decode configurations or sequence header.

4.2.3.4.1.2 Syntax

```
class AvstSequenceInfoBox extends Box('avss')
{
    AvstDecoderConfigurationRecord AvstConfig;
}
class BitRateBox extends Box('btrt')
{
    unsigned int(32) maxBitRate;
    unsigned int(32) avgBitRate;
}
class AvstSampleEntry() extends VisualSampleEntry ('avst'){
    AvstSequenceInfoBox SeqInfo;
    BitRateBox(); // optional
}
```

4.2.3.4.1.3 Semantics

Compressorname in the base class VisualSampleEntry indicates the name of the compressor used with the value “\016AVS2-P2 Coding” being recommended (\016 is 14, the length of the string as a byte).

maxBitrate gives the maximum rate in bits/second over any window of one second.

avgBitrate gives the average rate in bits/second over the entire presentation.

4.2.3.4.2 Sample format

4.2.3.4.2.1 Syntax

```
class AvstSample
{
    unsigned int PictureLength = sample_size;
    for (i=0; i<PictureLength;)
    {
        unsigned int((AvstDecoderConfigurationRecord.lengthSizeMinusOne+1)*8)
        NALUnitLength;
        bit(NALUnitLength * 8) NALUnit;
        i += (AvstDecoderConfigurationRecord.lengthSizeMinusOne+1) +
        NALUnitLength;
    }
}
```

4.2.3.4.2.2 Semantics

NALUnitLength indicates the size of a NAL unit measured in bytes, not including the length field itself.

NALUnit contains a single NAL unit.

4.2.3.5 IEEE 1857.10 video stream definition

4.2.3.5.1 Sample entry

4.2.3.5.1.1 Definition

Box types: 'avs3'

Container: Sample table box ('stbl')

Mandatory: An 'avs3' sample entry is mandatory in at least one track among the tracks carrying an IEEE 1857.10 video bitstream

Quantity: One or more sample entries may be present

Each sample entry of an IEEE 1857.10 video track shall be an IEEE 1857.10 video sample entry that has sample entry type equal to 'avs3' as specified in this subclause or 'lav3' as specified in 4.2.3.5.2. The following restrictions are set for IEEE 1857.10 video tracks:

- Each sample entry of an AVS3 track containing a conforming IEEE 1857.10 video bitstream as specified in IEEE Std 1857.10 shall contain an Avs3ConfigurationBox.
- Each sample entry of an IEEE 1857.10 video track containing the lowest temporal layer shall contain an Avs3ConfigurationBox, and an LAvs3ConfigurationBox.
- Each sample entry of an IEEE 1857.10 video track that does not contain the lowest temporal layer shall contain an LAvs3ConfigurationBox.

4.2.3.5.1.2 Syntax

```
class Avs3ConfigurationBox extends Box('av3c',version = 0) {  
    Avs3DecoderConfigurationRecord () Avs3Config;  
}  
class Avs3SampleEntry() extends VisualSampleEntry ('avs3') {  
    Avs3ConfigurationBox config;  
    LAvs3ConfigurationBox (); // optional    MPEG4ExtensionDescriptorsBox (); // optional  
}
```

4.2.3.5.1.3 Semantics

Avs3ConfigurationBox provides the decoding configuration information for IEEE 1857.10 video bitstream.

LAvs3ConfigurationBox provides the decoding configuration information for IEEE 1857.10 video bitstream that supports temporal layering.

MPEG4ExtensionDescriptorsBox is defined in ISO/IEC 14496-12:2021, section 5.4.2.

Compressorname in the base class VisualSampleEntry indicates the name of the compressor used with the value "\013AVS3 Coding" being recommended (\013 is 11, the length of the string in bytes).