SCTE | STANDARDS

Digital Video Subcommittee

AMERICAN NATIONAL STANDARD

ANSI/SCTE 242-3 2022

Next Generation Audio Coding Constraints for Cable Systems: Part 3 – MPEG-H Audio Coding Constraints

NOTICE

The Society of Cable Telecommunications Engineers (SCTE) Standards and Operational Practices (hereafter called "documents") are intended to serve the public interest by providing specifications, test methods and procedures that promote uniformity of product, interoperability, interchangeability, best practices, and the long term reliability of broadband communications facilities. These documents shall not in any way preclude any member or non-member of SCTE from manufacturing or selling products not conforming to such documents, nor shall the existence of such standards preclude their voluntary use by those other than SCTE members.

SCTE assumes no obligations or liability whatsoever to any party who may adopt the documents. Such adopting party assumes all risks associated with adoption of these documents and accepts full responsibility for any damage and/or claims arising from the adoption of such documents.

NOTE: The user's attention is called to the possibility that compliance with this document may require the use of an invention covered by patent rights. By publication of this document, no position is taken with respect to the validity of any such claim(s) or of any patent rights in connection therewith. If a patent holder has filed a statement of willingness to grant a license under these rights on reasonable and nondiscriminatory terms and conditions to applicants desiring to obtain such a license, then details may be obtained from the standards developer. SCTE shall not be responsible for identifying patents for which a license may be required or for conducting inquiries into the legal validity or scope of those patents that are brought to its attention.

Patent holders who believe that they hold patents which are essential to the implementation of this document have been requested to provide information about those patents and any related licensing terms and conditions. Any such declarations made before or after publication of this document are available on the SCTE web site at https://scte.org.

140 Philips Road
Exton, PA 19341

Document Types and Tags

Document Type: Specification

Document Tags:

☐ Test or Measurement	☐ Checklist	☒ Facility
☐ Architecture or Framework	☐ Metric	☐ Access Network
☒ Procedure, Process or Method	☐ Cloud	☐ Customer Premises

Document Release History

Release	Date
SCTE 242-3 2017	*10/10/2017*
SCTE 242-3 2022	*October 2022*

Note: Standards that are released multiple times in the same year use: a, b, c, etc. to indicate normative balloted updates and/or r1, r2, r3, etc. to indicate editorial changes to a released document after the year.

Table of Contents

List of Tables

1. Introduction

1.1. Executive Summary

This document is part of a suite documenting coding constraints of Next Generation Audio (NGA) systems for cable television (see [SCTE 242-1]).

1.2. Scope

This document is part of a suite documenting coding constraints of Next Generation Audio (NGA) systems for cable television. In conjunction with Part 1 of this standard [SCTE 242-1], it defines the coding constraints on MPEG-H Audio system for cable television.

The carriage of the streams described in this specification is defined in [SCTE 243-3] in conjunction with [SCTE 243-1].

1.3. Benefits

The MPEG-H Audio system ([ISO 23008-3] and [ISO 23003-4]) is a highly efficient audio system offering advanced methods for coding of channel-based content, coding of object-based content, and coding of scene-based content (using Higher Order Ambisonics (HOA) as a sound-field representation). An MPEG-H Audio encoded program might consist of a flexible combination of any of the audio program elements that are defined in Section 6 of this document:

- Channels (i.e., signals for specific loudspeaker positions),
- Objects (i.e., signals with position information) and
- Higher Order Ambisonics, HOA (i.e., sound field signals).

Other benefits of MPEG-H include a rich set of metadata that define Audio Scenes (also known as Audio Programs), Audio Objects, as well as elements that enable interactivity as well as personalization and accessibility options. The MPEG-H Audio system includes advanced tools for loudness and dynamic range control and enables seamless configuration changes in a broadcast environment as well as more advanced features available in ATSC 3.0.

This NGA system is not compatible with the audio system used in [SCTE 54]-era service.

2. Normative References

The following documents contain provisions which, through reference in this text, constitute provisions of this document. The editions indicated were valid at the time of subcommittee approval. All documents are subject to revision and, while parties to any agreement based on this document are encouraged to investigate the possibility of applying the most recent editions of the documents listed below, they are reminded that newer editions of those documents might not be compatible with the referenced version.

2.1. SCTE References

[SCTE 242-1] SCTE 242-1 202x, Next Generation Audio Coding Constraints for Cable Systems: Part 1 – Introduction and Common Constraints

2.2. Standards from Other Organizations

[A342-1] ATSC A/342-2021, Part 1: Audio Common Elements

[A342-3] ATSC A/342-2021, Part 3: MPEG-H System

[ISO 23008-3] ISO/IEC 23008-3:2019: Information technology -- High efficiency coding and media delivery in heterogeneous environments – Part 3: 3D audio, ISO/IEC 23008-3:2019/Amendment 1:2019, ISO/IEC 23008-3:2019/Amendment 2:2020.

[ISO 23003-4] ISO/IEC 23003-4:2020: MPEG audio technologies – Part 4: Dynamic Range Control

2.3. Other Published Materials

No normative references are applicable.

3. Informative References

The following documents might provide valuable information to the reader but are not required when complying with this document.

3.1. SCTE References

[SCTE 243-3] SCTE 243-3 202x, Next Generation Audio Carriage Constraints for Cable Systems: Part 3 – Carriage of MPEG-H Audio

[SCTE 243-1] SCTE 243-1 202x, Next Generation Audio Carriage Constraints for Cable Systems: Part 1 – Common Transport Signaling

[SCTE 54] SCTE 54 2020, Digital Video Service Multiplex and Transport Subsystem Standard for Cable Television

3.2. Standards from Other Organizations

[A85] ATSC: “Techniques for Establishing and Maintaining Audio Loudness for Digital Television,” Doc. A/85:2013, Advanced Television Systems Committee, Washington, D.C., 12 March 2013

[BS 1770-4] ITU: “Algorithms to measure audio programme loudness and true-peak audio level,” Recommendation ITU-R BS.1770-4, International Telecommunications Union, Geneva, Switzerland, 2015

3.3. Other Published Materials

No informative references are applicable.

4. Compliance Notation

shall	This word or the adjective "***required***" means that the item is an absolute requirement of this document.
shall not	This phrase means that the item is an absolute prohibition of this document.
forbidden	This word means the value specified ***shall*** never be used.
should	This word or the adjective "*recommended*" means that there *may* exist valid reasons in particular circumstances to ignore this item, but the full implications *should* be understood and the case carefully weighed before choosing a different course.
should not	This phrase means that there *may* exist valid reasons in particular circumstances when the listed behavior is acceptable or even useful, but the full implications *should* be understood and the case carefully weighed before implementing any behavior described with this label.
may	This word or the adjective "*optional*" indicate a course of action permissible within the limits of the document.
deprecated	Use is permissible for legacy purposes only. Deprecated features *may* be removed from future versions of this document. Implementations *should* avoid use of deprecated features.

5. Abbreviations and Definitions

5.1. Abbreviations

DRC	Dynamic Range Control
HOA	Higher Order Ambisonics
LC	Low Complexity
MAE	MPEG-H Metadata Audio Element
MHAS	MPEG-H Audio Stream
RAP	Random Access Point

5.2. Definitions

This document uses the terminology defined in Part 1 of this standard [SCTE 242-1], and the mapping of the ATSC 3.0 Audio glossary terms to the MPEG-H audio alternative terms as defined in ATSC [A342-1] Clause 4.

6. MPEG-H Audio Coding Constraints

6.1. Introduction

The MPEG-H Audio system offers methods for coding of channel-based content, coding of object-based content, and coding of scene-based content (using Higher Order Ambisonics [HOA] as a sound-field representation). An MPEG-H Audio encoded program *may* consist of a flexible combination of any of the audio program elements that are defined in Part 1 of this standard ([SCTE 242-1]), subclause 6.2.2, namely:

- Channels (i.e., signals for specific loudspeaker positions),
- Objects (i.e., signals with position information) and
- Higher Order Ambisonics, HOA (i.e., sound field signals).

The MPEG-H Audio system is a Next Generation Audio (NGA) system standardized in ATSC 3.0. A complete overview of the MPEG-H Audio system and its main features is given in ATSC [A342-3] Clause 4 and ATSC [A342-1] Clause 5.

6.1.1. Profiles and Levels

Table 1 is provided here for information and it shows which are the limitations of each level of the Low Complexity Profile and examples of configurations supporting the maximum values for each Level.

Table 1 - Levels and their corresponding restrictions for the MPEG-H Audio LC Profile (Informative)

Level	Max. Sampling rate	Max. no. of core ch. in compressed data stream	Max. no. of decoder processed core ch.	Max. no. of loudspeaker output ch.	Example of max. loudspeaker configuration	Max. no. of decoded objects	Example of a max. Config C+O	Max. HOA order	Example of max. HOA order + O
1	48000	10	5	2	2.0	5	2 ch. + 3 static obj.	2	2nd order + 3 static obj.
2	48000	18	9	8	7.1	9	6 ch. + 3 static obj.	4	4th order + 3 static obj.
3	48000	32	16	12	11.1	16	12 ch. + 4 obj.	6	6th order + 4 obj.
NOTE – In this context "static objects" are understood as channel-based signals without accompanying OAM data which are not also associated to a channel bed									

6.2. Encoding Constraints

Bitstreams produced by the encoder ***shall*** comply with the MPEG-H Audio Low Complexity (LC) Profile as defined in [ISO 23008-3] Clause 4. The encoder ***shall*** use MPEG-H Audio LC Profile Level 1, Level 2 or Level 3. If the audio signals are encoded according to the Baseline (BL) Profile restrictions specified in [ISO 23008-3] Clause 4.8.2.6, the CompatibleProfileLevelSet() config extension element specified in [ISO 23008-3] Clause 4.8.2.7 ***shall*** be present. The mpegh3daProfileLevelIndication and CompatibleSetIndication ***shall*** be as set according to Table P.1 in [ISO 23008-3].

All constraints specified in Part 1 of this standard ([SCTE 242-1]), Clause 7 ***shall*** apply.

The MPEG-H Audio elementary stream data will be encapsulated into MPEG-H Audio Stream (MHAS) packets according to [ISO 23008-3] Clause 14.

6.3. Expectations for decoders

It is expected that decoders will support MPEG-H Audio LC Profile Level 1, Level 2 and Level 3 as defined in [ISO 23008-3] Clause 4. Operators *should* therefore consider the capabilities of decoders that *may* be available to their customers in deciding whether to transmit bitstreams coded in accordance with Level 1, Level 2 or Level 3. It is also expected that all decoders will support MHAS transport format.

7. Metadata

7.1. Metadata Audio Elements

MPEG-H Audio uses a set of static metadata, the "Metadata Audio Elements" (MAE), to define an "Audio Scene". An Audio Scene represents an Audio Program as defined in ATSC [A342-1] Clause 4.

If one Audio Program contains Audio Preselection description, the Audio Program ***shall*** contain exactly one default Audio Preselection, i.e., the Audio Preselection containing the main audio to be decoded in the absence of any user preference data or any other system automatic selection information. This means that, if the Audio Program contains Audio Preselection description, exactly one group preset ***shall*** have the **mae_groupPresetID** field set to 0**.**

The number of Audio Preselection ***shall*** be equal with 31 or less, i.e., the **mae_numGroupPresets** field ***shall*** be set to a value between 0 and 31. If the Audio Program does not contain Audio Preselection description the **mae_numGroupPresets** field ***shall*** be set 0.

7.2. Loudness and Dynamic Range Control

The MPEG-H Audio system includes advanced tools for loudness and dynamic range control inherited from MPEG-D DRC [ISO 23003-4]. MPEG-D DRC defines a comprehensive and flexible metadata format that is compliant to worldwide loudness regulations including those based on ATSC [A85].

Loudness metadata embedded in MPEG-H Audio elementary streams ***shall*** comply with the constraints defined in ATSC [A342-3] subclause 5.3.1. Methods to measure loudness dependent on content type and presence of an Anchor Element are described in ATSC [A85]. The loudness measurement algorithm specified in ITU-R [BS 1770-4] *should* be applied.

8. Random Access

Random access and immediate play-out is possible at every frame that utilizes the AudioPreRoll() structure as specified [ISO 23008-3] subclause 5.5.6.

The MPEG-H Audio data contained in the mpegh3daFrame() structure ***shall*** follow the rules for random access points (RAPs) as defined in [ISO 23008-3] subclause 5.7.

9. Configuration Change and Audio/Video Alignment

MPEG-H Audio enables seamless configuration changes in a broadcast environment. A configuration change takes place in an audio stream when the content setup or the Audio Scene Information changes (e.g., when changes occur in the channel layout, the number of objects, etc.). Even though configuration changes usually happen at program boundaries, they are not restricted to that case and they *may* occur at any time within a program.

10. Multi-Stream delivery

The multi-stream-enabled MPEG-H Audio system is capable of handling Audio Program Components delivered in several different elementary streams (e.g., one MHAS stream containing one complete audio main, and one or more additional MHAS streams, containing different languages and audio description).

The main MHAS stream (containing one complete audio main or the Audio Program Components corresponding to the default Audio Preselection) and the additional MHAS streams (containing the Audio Program Components corresponding to several other Audio Preselections) can be carried:

- within a single MPEG-H Audio elementary stream, or
- as separate MPEG-H Audio elementary streams.

The MAE information allows the MPEG-H Audio Decoder to correctly decode several MHAS streams. The MHAS streams can be provided directly to the MPEG-H Audio Decoder. Alternatively, the MHAS streams can be first merged into one single MHAS stream, which is provided to the MPEG-H Audio Decoder.