

SMPTE ENGINEERING GUIDELINE

System Implementation of CEA-708 and CEA-608 Closed Captioning and Program-Related Data



Page 1 of 32 pages

Table of Contents	Page
Foreword	3
Intellectual Property	3
Introduction	3
1 Scope	4
2 Conformance Notation	4
3 Terms and Acronyms	5
3.1 Acronyms and Abbreviations	5
3.2 Terms	5
4 Closed Captioning Standards	6
4.1 Relevant CEA Standards	6
4.2 Relevant SMPTE Standards	7
4.3 Relevant ATSC Standards	7
4.4 Relevant SCTE Standards	8
4.5 Relevant DVB/ETSI Standards	9
5 System Configurations	9
6 Caption Creation and Encoding	9
6.1 Captioning Process	9
6.2 CEA-608 Caption Encoding	10
6.3 CEA-708 Caption Encoding	11
6.4 CEA-708 Caption Encoding Feeding Direct to the MPEG Encoder	14
6.5 CEA-708 Caption Translation within the MPEG Encoder Using CEA-608 Captions from Line 21	14
6.6 Required Level of Implementation	15
6.7 Adding Additional DTV Captions after Initial Caption Authoring	15
7 Recording, Processing and Routing	15
7.1 Recording	16
7.2 VANC Data Bridge	16
7.3 Video Processing	17
7.4 Switching Considerations	17
8 Contribution and Distribution	17
8.1 Tape Interchange	18
8.2 File-Based Delivery	18
8.3 Contribution/Distribution Compression Systems	18

9	Emission	19
9.1	NTSC Output	20
9.2	ATSC Station Output	20
9.3	Captioning Inputs to an MPEG Emission Encoder	21
9.4	Captioning in AVC Bitstreams	22
9.5	Captioning for ATSC Mobile DTV	22
9.6	Local Live DTV Captions	22
9.7	NTSC – DTV (Up-Conversion)	23
9.8	DTV – NTSC (Down-Conversion)	23
9.9	Completeness of CEA-608 Captions, Text, and Program data	23
9.10	Consistency of CEA-708 and CEA-608 Caption Services	24
9.11	Consistency of XDS Data and PSIP Data	24
9.12	ANSI/SCTE 127 Data	24
9.13	Server-Based Captioning	24
9.14	Caption Service Descriptor	25
10	Monitoring	26
10.1	608 Caption Monitoring	26
10.2	708 Caption Monitoring	26
Annex A	Bibliography (Informative)	28
Annex B	Descriptor Enquiry Message Format (Informative)	30
Annex C	Field-Marked Message Format for CEA-608 Data (Informative)	32

Foreword

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

SMPTE Engineering Documents are drafted in accordance with the rules given in Part XIII of its Administrative Practices. This SMPTE Engineering Document was prepared by Technology Committee 10E.

Intellectual Property

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

Specifications for closed captioning for NTSC analog television broadcasting¹ are set out in the CEA-608 standard. Production, distribution and insertion of captioning information to meet these requirements have been implemented for many years using a combination of standards-based and accepted-practice techniques.

Specifications for closed captioning for ATSC digital television (DTV) in the USA are set out in the CEA-708 standard. The term DTV in CEA-708 includes and applies to HDTV (high definition digital television) and SDTV (standard definition digital television) that use the digital data stream specified in ATSC A/53. CEA-708 captions are also specified for use in ATSC standard A/72 for AVC (Advanced Video Coding), in SCTE cable systems for both AVC (ANSI/SCTE 128) and SMPTE VC-1 (SCTE 157), as well as in ATSC A/153 for Mobile/DTV.

During the transition from analog to digital broadcasting, analog NTSC and DTV transmissions will coexist, as will analog and digital cable distribution systems. After the closure of NTSC transmissions, analog and digital television production and distribution systems will possibly coexist in many facilities and networks, frequently with both standard definition (SD) and high definition (HD) video formats. This situation creates added complexity for closed captioning as well as other aspects of system implementation. In addition, some multichannel video program distribution systems will continue to distribute NTSC analog signals and in some cases such signals are derived by conversion from ATSC DTV signals (see ATSC A/79 for more information). In other cases, NTSC signals will be derived from ATSC DTV signals by a consumer digital converter box.

CEA-708 specifies standards for encoding and decoding DTV closed captions (DTVCC) but provides little guidance on system design and interfacing for DTVCC in broadcast systems.

¹ NTSC broadcasting from full-power NTSC stations in the USA ended in June 2009 but NTSC transmissions from low-power stations and translators will continue for some time after that date. NTSC broadcasts also continue in other countries.

1 Scope

The primary purpose of this guideline is to provide guidance for system implementation of closed captioning for DTV as defined in CEA-708, concentrating on different techniques that are required for DTVCC implementation when compared to CEA-608 captioning systems for NTSC. It also provides guidance on distribution of certain program-related data carried in the CEA-608 data construct. It identifies the relevant standards that exist for different parts of the system and, where possible, indicates guidelines for areas that are not standardized.

This guideline describes methods that can be used to support the various links in the chain from caption creation through to emission. Consideration is given to systems that support captioning for SD and HD program material, which can originate and/or be distributed in standard formats as used in broadcasting and related facilities. While CEA-708 refers specifically to DTV implemented for mainstream ATSC broadcast systems, techniques described herein are applicable for programming for other DTV environments such as ATSC Mobile DTV and cable or satellite distribution.

In this guideline, transport for distribution of DTVCC within the broadcast plant is based primarily on the use of SMPTE 334-1, which define a method of embedding DTVCC and other data services in the vertical ancillary (VANC) data space of video signals conforming with SMPTE 292 high-definition serial digital interface (HD-SDI) or SMPTE 259M serial digital interface (SDI). HD-SDI in this guideline refers to bit-serial component 1080 or 720 line television signals. SDI in this guideline refers to bit-serial component 525-line television signals. The carriage of DTVCC over serial data links in accordance with SMPTE 333 or RP 2007 is also included, together with arrangements for distributing caption service information. Reference is made to alternative server-based methods of distributing and managing closed captioning.

A method for carriage of DTVCC caption data in an AES3 data stream is specified in SMPTE 337M. Further details of that arrangement are not described here.

2 Conformance Notation

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

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

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

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

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

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

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

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

3 Terms and Acronyms

3.1 Acronyms and Abbreviations

AES3: AES standard for serial digital audio

ATSC: Advanced Television Systems Committee

CDP: Caption Distribution Packet

CEA: Consumer Electronics Association

CSD: Caption Service Descriptor

DTVCC: DTV Closed Captions

DTV: Digital Television

EIT: Event Information Table

NTSC: National Television System Committee

PSIP: Program and System Information Protocol

HD-SDI: High Definition Serial Digital Interface (compliant with SMPTE 292)

SCTE: Society of Cable Telecommunications Engineers

SDI: Serial Digital Interface (compliant with SMPTE 259M)

VANC: Vertical ANCillary (data space)

3.2 Terms

Caption Encoder: A device that encodes closed caption data and inserts it into a video signal. The term is used both for equipment inserting CEA-608 captions (as line 21 data) and for equipment inserting CEA-708 captions (usually in the form of CDPs carried in VANC). This term must not be confused with an emission or MPEG encoder.

Caption Server: (1) a computer that sends data from a caption intentions file (produced by a caption author) to a caption encoder. (2) In SMPTE 333 the term is used to refer to a device that delivers DTV closed caption formatted data to an MPEG Encoder for insertion in a video elementary stream.

Emission encoder: A compression encoder processing video, audio, and associated data services. As used in this document the video is compressed as MPEG-2 and the audio as AC-3. The emission encoder outputs an MPEG-2 Transport Stream per ATSC A/53.

MPEG Encoder: Device which compresses video, audio, and associated data using (historically) MPEG-2. In this document, this term is understood to be generalized to include additional compression standards, including AVC and SMPTE VC-1.

Uplink encoder: A compression encoder processing video, audio, and associated data services for a Contribution/Distribution system. Video compression might be MPEG-2, AVC, VC-1 or any other suitable video compression codec. Audio compression might be AC-3, MPEG-1 Layer II or any other suitable audio compression codec.

Care must be taken in discussing closed captioning with program and captioning providers to avoid using the ambiguous term “subtitles”. Use of this term can lead to misunderstanding, particularly with international communications, as the term “subtitles” is well defined in other contexts which do not match the use of CEA-708 (or CEA-608) captions. In these circumstances, the term “open captions” might be used in preference to “subtitles” to make clear the distinction from closed captions.

4 Closed Captioning Standards

The role of the primary standards relating to captioning can be summarized as follows:

4.1 Relevant CEA Standards

4.1.1 CEA-608

CEA-608 defines the coding of data, including captioning, which is carried in line 21 (and line 284) of an analog NTSC video signal or (as a digitized analog waveform) in line 21 of an SDI digital video signal. It includes the specifications and techniques for encoding, insertion, extraction, re-encoding, transmission, reception, and display of caption data and defines non-caption data that can be carried, including Text Mode data service and Extended Data Service (XDS), which can include information about the program, content advisory, and various other data.

CEA-608 refers only to encoding data on line 21, but the data structures defined can be used to carry information on other lines in the video signal, usually in the vertical blanking interval. Such alternative uses are not covered in this guideline.

In this guideline, captions generated in accordance with the CEA-608 standard are referred to as CEA-608 captions; they are sometimes referred to simply as “608” captions.

4.1.2 CEA-708

CEA-708 defines the coding of DTVCC as they are delivered in an ATSC A/53 emission bitstream, and is applicable equally to HD and SD video formats, and to terrestrial broadcasting and cable and satellite distribution. It includes a description of the specific data packets and structures of DTVCC. The standard provides a specification of how caption information is to be coded and processed, minimum implementation recommendations for DTV closed caption decoders, and recommended practices for DTV closed caption encoder and decoder manufacturers.

Caption data delivered in an ATSC bitstream includes the data required for use by a DTV integrated receiver or set-top box to display DTV captions, often referred to as “708” data. It also includes CEA-608 data required for use by DTV set-top boxes so that line 21 data can be inserted in an NTSC output (if provided) to feed an analog TV set. The CEA-608 data also enables line 21 captioning and data insertion for any other equipment that needs to produce an NTSC output from a down-converted HD stream. CEA-708 indicates that DTV receivers are allowed to use CEA-608 caption data when CEA-708 data is not available for providing closed captioning—note this is not mandatory and is not implemented in many receivers; DTV receivers must use CEA-708 data as the default DTV closed caption format. The CEA-708 document outlines a TV captioning “food chain” from authoring to emission but refers to this SMPTE engineering guideline for information on station operations and caption handling.

In this guideline, captions generated in accordance with the CEA-708 standard are referred to as CEA-708 captions or as DTVCC; they are sometimes referred to simply as “708” captions.

4.2 Relevant SMPTE Standards

There are several relevant SMPTE standards which should be noted.

4.2.1 SMPTE 334-2

SMPTE 334-2 defines a Caption Distribution Packet (CDP) consisting of a specific sequence of bytes that can hold: the CEA-708 DTV caption data, CEA-608 caption data, caption service information, and SMPTE 12M-1 time code. The CDP is the basic unit of data that is transported through the professional portion of a DTV caption distribution chain. As such it is central to several of the methods discussed in this guideline.

Caption service information in the CDP includes data needed to form the caption service descriptor as defined in ATSC A/65 and used in the ATSC transport stream to announce the presence and format of captions being carried. The optional time code can be used in resynchronizing captions with video if the CDP is carried separately from the video signal.

4.2.2 SMPTE 334-1

SMPTE 334-1 defines a method of placing CDPs into the VANC space of either a standard definition or a high definition digital video signal.

4.2.3 SMPTE 333 and SMPTE RP 2007

SMPTE 333 defines a method of carrying CDPs from a caption encoder to an MPEG encoder via a serial interface (EIA/TIA-232). The protocol defined uses software “handshaking” for flow control and synchronization between the two devices.

There is an alternative protocol described in SMPTE RP2007 which does not use flow control, so the caption data is “pushed” from the caption encoder to the MPEG encoder.

Both methods are more fully discussed later in this document.

4.3 Relevant ATSC Standards

The ATSC standards provide for the compression, transport, and transmission of digital television signals via terrestrial means. There are a number of different ATSC standards which have relevance in a discussion of DTVCC carriage. These are discussed in the following sections.

4.3.1 ATSC A/53

ATSC A/53 is a seven-part document that specifies the ATSC digital television system. A/53 Part 4 specifies the MPEG-2 video compression which is used, along with the syntax and semantics for the encoding and transport of caption data contents as specified in CEA-708.

4.3.2 ATSC A/65

ATSC A/65 defines information in the Program and System Information Protocol (PSIP) that describes the contents of an ATSC broadcast. It defines the caption service descriptor (CSD) carried in the transmitted bitstream to signal the service number and language of digital (708) captions, if present, and indicate the presence of analog (608) captions. This data can be used by the DTV receiver to display information about captioning in an electronic program guide, and will possibly be needed to properly decode and display the closed captions. A/65 states that for terrestrial broadcast transport streams the CSD shall be present in the EIT (Event Information Table) and may be present in the PMT (Program Map Table). If present in the PMT, the data must match the current EIT-0 for the corresponding virtual channel. For transport streams delivered on

cable the CSD must be present in the PMT for captioned programs, and if an EIT is sent, must also be present in the EIT.

4.3.3 ATSC A/72

The use of AVC in the ATSC digital television system is documented in ATSC A/72. A/72 Part 1 specifies that closed captions shall be transported in accordance with A/53 Part 4, with syntax and semantics as detailed by A/72 Part 1, and with the caption data contents as specified in CEA-708.

4.3.4 ATSC A/153

ATSC/153 ATSC is an eight-part document that specifies the ATSC Mobile DTV system. A/153 Part 7, AVC and SVC Video System Characteristics, specifies AVC video compression and uses the same method for carrying closed captions in the AVC bitstream as A/72 Part 1, except that variable bit rates, not to exceed 9600 bits per second, are to be used for the closed captioning payload. That is, packing bytes are not to be used, and when captions are not present, there is no bandwidth allocation. This is a permissible difference from the earlier versions of CEA-708, which required allocation of 9600 bits per second for the closed caption payload data for all DTV system bitstreams.

Part 3 of the Mobile DTV standard, Service Multiplex and Transport Subsystem Characteristics, includes the specification for signaling metadata describing the contents of a Mobile DTV broadcast. It specifies that the caption service descriptor as defined in ATSC A/65C shall be used to provide closed captioning information if one or more closed captioning services are provided in the Mobile DTV service.

4.4 Relevant SCTE Standards

Most ATSC signals are ultimately carried on cable systems, typically governed by standards adopted by the SCTE. There are a number of different SCTE standards which have relevance in a discussion of DTVCC carriage as well as CEA-608 carriage. These are discussed in the following sections.

4.4.1 ANSI/SCTE 40

ANSI/SCTE 40 defines the overall system characteristics for cable television, including RF. It specifies that for digital television programs, when closed caption data is provided, it is to be provided in accordance with one or both of two closed caption data formats as specified by ANSI/SCTE 43.

4.4.2 ANSI/SCTE 43

ANSI/SCTE 43 defines the digital video system characteristics for cable television. It states that advanced DTV closed captions, when present, are to be encoded in accordance with CEA-708 and shall be transported in accordance with ATSC A/53. It also specifies that CEA-608 Line 21 captions, when present, are to be transported in accordance with ANSI/SCTE 20 or ATSC A/53 or both (when both are present, it defines the order in which they must occur).

4.4.3 ANSI/SCTE 54

ANSI/SCTE 54 defines the digital video service multiplex and transport standard for cable television. It states that when caption services are delivered within the picture data construct defined in ATSC A/53, the caption service descriptor shall be present in the PMT and in the optional EIT.

4.4.4 ANSI/SCTE 20

ANSI/SCTE 20 defines a method for carrying CEA-608 Line 21 captions in a similar manner to that specified by A/53 for carrying CEA-708 captions. This method was developed prior to the publication of the original

A/53 and is similar to numerous proprietary mechanisms used by various manufacturers for the same purpose. ANSI/SCTE 20 can also be used to carry certain VBI data structures.

4.4.5 ANSI/SCTE 21

ANSI/SCTE 21 defines a standard for the carriage of CEA-608 Line 21 captions along with Vertical Blanking Interval (VBI) services in MPEG-2 compliant bitstreams. The approach extends the data structure defined in the ATSC A/53 Digital Television Standard, and is compatible with that method.

4.4.6 ANSI/SCTE 128

The use of AVC in SCTE digital cable systems is documented in ANSI/SCTE 128, which specifies that closed captions shall be transported in accordance with A/53 Part 4, with syntax and semantics as detailed by ANSI/SCTE 128, and with the caption data contents as specified in CEA-708.

Note that the syntax and semantics as detailed by ANSI/SCTE 128 are identical to those detailed in ATSC A/72 Part 1.

4.4.7 ANSI/SCTE 157

The use of SMPTE VC-1 in SCTE digital cable systems is documented in ANSI/SCTE 157, which specifies that closed captions shall be transported in accordance with A/53 Part 4, with syntax and semantics as detailed by ANSI/SCTE 157, and with the caption data contents as specified in CEA-708.

4.5 Relevant DVB/ETSI Standards

Digital Video Broadcasting (DVB) added CEA-708 closed captions to Annex B of TS 101 154 in version 1.8.1 (2007). It documents CEA-708 caption data carriage by MPEG-2, AVC, and VC-1 in identical manner in each case with ATSC A/53 Part 4, ATSC A/72 Part 1, and ANSI/SCTE 157, respectively.

5 System Configurations

Overall systems for implementing closed captions will vary considerably from user to user, and arrangements possibly will change dynamically depending on program requirements. No single diagram can illustrate the configuration that will apply to all users. However, certain system elements relating to captioning will be required in most systems. These comprise the functional blocks of creation, encoding, recording, routing and processing within a plant, contribution and distribution, and emission.

The following sections provide guidelines for implementing each of these functional blocks for captioning in a DTV environment. The drawings are intended to be indicative of possible arrangements and users are cautioned not to take them as literal system diagrams or absolute designs since product features in this area will continue to evolve.

6 Caption Creation and Encoding

6.1 Captioning Process

The process for creating closed captions for recorded material typically consists of three steps:

1. A caption author views the video/audio program and, using a software-based authoring system running on a computer with a time code reader, creates a file that associates text strings and their display attributes with SMPTE 12M-1 time codes that define the insertion point in the program. This file is known as a "caption intentions file";

2. When the captions are to be inserted into the video signal a computer running captioning software reads the caption intentions file and, in response to a time code input, at the specified times sends commands containing data from the intentions file to a caption encoder at the point of insertion. The computer is often referred to as a caption server (this is not the same as the caption data server referred to in SMPTE 333, see 9.3);
3. In response to the commands received from the computer, either the caption encoder embeds caption data into the video signal of the related program material, in a suitable format for the signal type as discussed in 6.2, 6.3, and 6.4, or the caption server sends the caption data to the video encoder via SMPTE 333 as discussed in 9.3 method 3. In both cases the response to the commands from the computer must be frame accurate.

6.1.1 Captioning for live programming

For live programming, only step 3 above applies. In that case data is fed from a specialized authoring workstation to the caption encoder, which immediately inserts the captions into the video stream ready for transmission. The authoring workstation can be connected directly to the caption encoder or can be remotely located, where it typically sends the captioning signals over telephone lines using a modem, or via the Internet.

6.2 CEA-608 Caption Encoding

608 caption encoding for recorded material is shown in Figure 1. The program video is typically analog NTSC or digital SDI. The data from the captioning computer is processed and inserted in line 21 of the video to produce a signal that complies with CEA-608 for NTSC transmission.

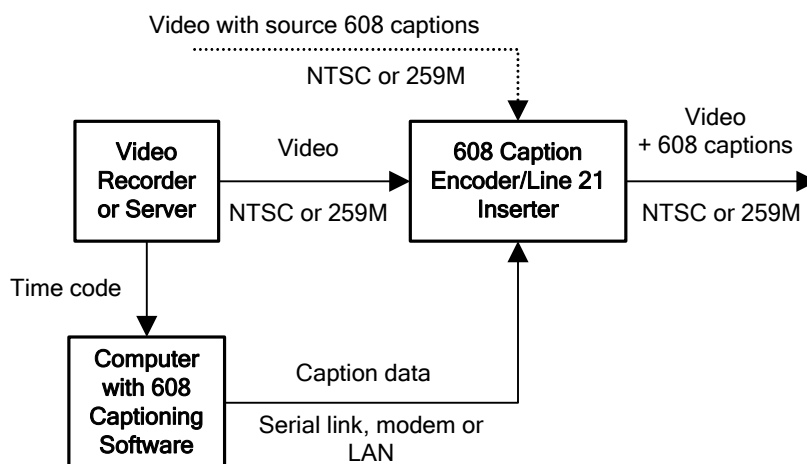


Figure 1 – Typical CEA-608 caption encoding

Points to note are:

- a. SMPTE 12M-1 time code is universally used.
- b. Various manufacturers have proprietary formats for caption intention files. Captioning computers and encoders can generally work with files generated by a wide variety of caption authoring software. Information on compatibility is available from the manufacturers.
- c. Captioning computer to caption encoder communications can be via serial port, modem, or local area network (LAN) connection. There are no published standards for captioning computer to caption encoder