SMPTE EG 43:2009
Revision of EG 43-2004

SMPTE ENGINEERING GUIDELINE

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

Table of Contents

Approved October 23, 2009

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 broadcasting[1] 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.

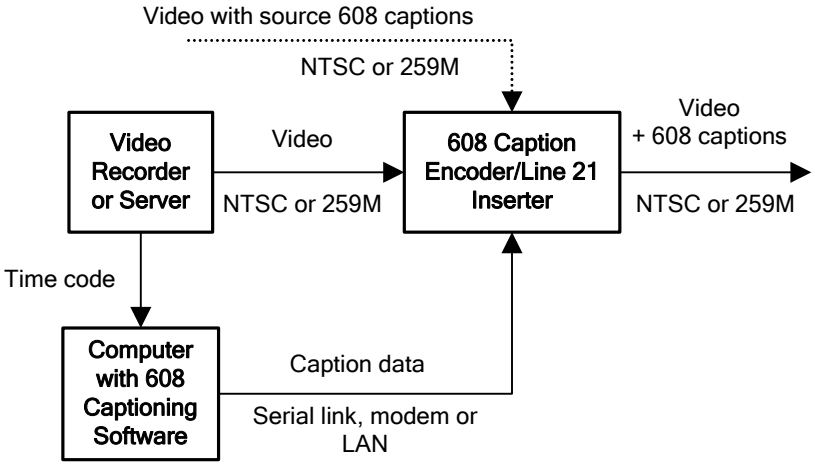

Figure 1 – Typical CEA-608 caption encoding

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

communications. One format has become a de facto industry standard for CEA-608 closed captioning although some manufacturers have slightly different implementations. Information on interface details and compatibility is available from the manufacturers.

d. Most caption encoders can also take as a caption source line 21 data that has been extracted from another video signal source, shown dotted in Figure 1. In some cases this extracted caption data will be bridged from another unit, typically using a serial data link, not shown in Figure 1. Interface arrangements for this are manufacturer-specific and two examples of message formats are provided in Annex C.

e. CEA-608 defines only the analog waveform for the data to be carried on line 21. Caption encoders for SDI signals encode the equivalent digital values of the defined waveform, as determined by the manufacturer.

f. Caption encoding and insertion on line 21 is an essential requirement for CEA-608 captioning. Since line 21 is transported over most parts of an analog NTSC distribution chain as an integral part of the video signal, it is generally straightforward to insert, extract, or re-encode captions at any point where the video program is available as an analog NTSC or SDI signal. Note that DTV signals compressed in accordance with RP 202 do not include line 21.

g. Monitoring of CEA-608 caption encoding and content must be carried out at appropriate points in the distribution chain. This is discussed further in 10.1.

h. In fully-digital facilities, it is also possible that CEA-608 caption data could be carried in a VANC CEA-608 packet per SMPTE 334-1 rather than simply as digitized versions of a CEA-608 waveform in line 21. Such VANC data would need to be encoded on line 21 for NTSC transmission and at the time of writing this guideline this arrangement is not being implemented. See also 6.3 (g).

6.3 CEA-708 Caption Encoding

Where CEA-708 captions need to be carried in a video signal, as shown in Figure 2 the architecture for encoding is similar to the CEA-608 arrangement. The program video in this case is HD-SDI or SDI.

The data from the captioning computer is processed and inserted as caption distribution packets embedded in VANC packets in the video signal, in accordance with SMPTE 334-1 and SMPTE 334-2. The CDP is used as the payload because it can encapsulate the caption-related information needed for downstream processing and emission, including all or some of:

- CEA-708 captions
- CEA-608 caption data
- Caption service information to form the caption service descriptor
- Time code

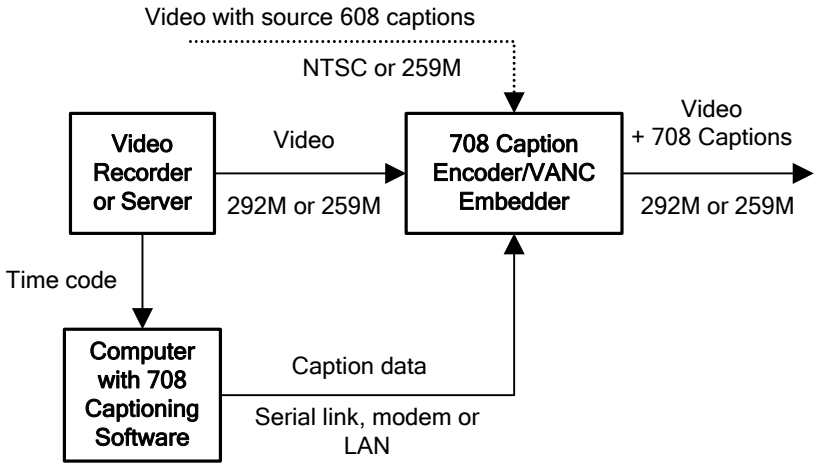

Figure 2 – Typical CEA-708 caption encoding

Points to note are:

a. This technique is independent of production format and equally applicable to captioning HD and SD program material.

b. The time code interface is still SMPTE 12M-1. The CDP can carry an (optional) time code derived from the original video source VITC or LTC time code. The presence of time code in the CDP makes it possible for downstream equipment to synchronize the captions and video if they flow through differing paths through the broadcast chain.

c. In some implementations the CEA-708 caption encoder/VANC embedder could have a built-in disk drive able to read a caption intentions file directly without an external computer.

d. At the time of writing this guideline work is in progress by US national captioning organizations to try to establish an agreed CEA-708 caption intentions format, but there is currently no such published protocol. Some CEA-708 caption encoders can accept data from files configured with data in native CEA-708 format, i.e., exactly as defined in CEA-708. The caption data provided to the encoder must also contain the associated CEA-608 captions. Information on compatibility is available from the manufacturers.

e. The format of the VANC CEA-708 closed captioning data packets is defined in SMPTE 334-1. The CDP payload format, containing the elements in the bullet list above, is defined in SMPTE 334-2.

f. The caption encoder can embed the CEA-708 captions, the CEA-608 caption data (also known as CEA-608 “compatibility bytes”) and the caption service information. Only a single instance of CEA-708 caption data per frame of video is allowed to be placed in VANC.

g. SMPTE 334-1 also defines separate VANC data packets for CEA-608 closed captioning data. That method of carrying CEA-608 data is additional to the one described above for encapsulation within the CEA-708 CDP. Caption systems being implemented at the time this guideline is written generally encapsulate CEA-608 data only within the CEA-708 CDP, and the dedicated CEA-608 VANC packet is typically not used.

h. SMPTE 334-1 also defines a separate VANC data packet for program description information required for insertion into PSIP tables. As set out in SMPTE RP 207, this packet could include the caption service descriptor. That method of carrying the CSD is additional to the one described above for encapsulation within the CEA-708 CDP. Caption systems being implemented at the time this guideline is written generally encapsulate the CSD data only within the CEA-708 CDP.

i. This method of DTV caption encoding can take place at any point where the video program is available as an HD-SDI or SDI signal. For CEA-708 captioning there are currently three modes of operation that depend on the native format of the caption creation tools, as described in 6.3.1, 6.3.2 and 6.3.3.

j. Currently manufactured CEA-708 caption equipment allows for CDPs distributed with a video signal (say from a network) to be modified if necessary at a downstream location (for example at a broadcast station) so that captions can be added or modified (see 6.7).

k. Monitoring of CEA-708 and associated CEA-608 caption encoding and content must be carried out at appropriate points in the distribution chain. This is discussed further in 10.2.

l. CEA-708 refers to SMPTE RP 218 for specification of the safe title area, which affects the location of captions on screen. For program production, RP 218 has been effectively replaced by SMPTE 2046-1, which specifies a larger safe title area than previously used. Caption authors need to be aware that although a larger area of the picture can now be used for non-caption titling, the constraints of RP218 for caption authoring and display, as specified in CEA-708, have not changed.

6.3.1 Native CEA-708 encoding

In this case, arrangements are as indicated in 6.3 above using data sent from the caption server to the caption encoder. The captions are authored for CEA-708 use, so all of the extensive display features provided by the CEA-708 standard are available. The interface between the caption server and the caption encoder must allow for the carriage of both CEA-708 and CEA-608 data. Information on compatibility is available from the manufacturers.

6.3.2 CEA-708 encoding based on translation of CEA-608 caption intentions

In this case, the captions are encoded in the CEA-708 DTVCC format, although they are authored and transported to the encoder in the format normally used for CEA-608 encoding. This will happen when the caption creator chooses to use CEA-608 creation tools for DTV programming. This is expected to be an interim solution until CEA-708 caption authoring is generally adopted.

Arrangements are as indicated in 6.3, with the following points to note:

a. Most CEA-708 caption encoders can accept data derived from CEA-608 caption intention files. See 6.2 (c) for caption server to caption encoder interface issues. Information on compatibility is available from the manufacturers.

b. The resulting DTV captions are limited to the authoring capabilities of CEA-608 and cannot exercise all features of a CEA-708 caption decoder. However it is possible for a viewer with a CEA-708 caption decoder to over-ride the transmitted settings and change the default font, size, color, and background.

6.3.3 CEA-708 encoding based on translation of decoded CEA-608 captions from line 21

In this case, the captions are generated in the CEA-708 DTVCC format, but are translated from CEA-608 legacy captions already encoded on line 21 of an incoming analog NTSC or SDI video signal. This method can be used when the program, including captions, was created primarily for NTSC transmission but is being encoded for DTV transmission, either as an SD program or with up-conversion to HD. For new programming this is expected to be an interim solution until native CEA-708 caption authoring is generally adopted, but could persist indefinitely for legacy NTSC material that already includes CEA-608 captions.

Arrangements are as indicated in 6.3, with the following points to note:

a. Instead of using caption data from a caption server, the CEA-708 caption encoder takes as a caption source line 21 data, which has been decoded, from an analog NTSC or digital SDI video signal. The

video input for this is shown dotted in Figure 2. With some manufacturers, the line 21 decoding is carried out with a separate piece of equipment (not shown in Figure 2), which then sends the data over a serial link to the CEA-708 caption encoder. Interface arrangements for this are manufacturer-specific and two examples of message formats used are provided in Annex C. See 9.7 for an application of this arrangement.

b. The resulting DTV captions are usually known as “translated CEA-708”, with limitations as noted in 6.3.2 (b). They are also referred to in the FCC 00-259-2000 Report and Order, “Closed Captioning Requirements for Digital Television Receivers” as “upconverted captions”. It is noted that some manufacturers and others use the term “transcoded” in this regard, but preferably that is used only to refer to the CEA-608 compatibility bytes, which are transcoded from CEA-608 caption sources.

c. Such “data bridging” must be done carefully to correct for latencies in related encoding and decoding cycles, and assumes that the program content of the source for the line 21 caption decoder and for the CEA-708 caption encoder is identical. It is further noted that, although CEA-708 makes some recommendations, there is no standard defining this “translation” (or “upconversion”) operation, so each manufacturer’s implementation is likely to be different.

6.3.4 Multiple caption services

Multiple caption services can be associated with a single DTV program. CEA-708 states that there shall be a maximum of 16 separate services present, with at most one CEA-608 datastream signaled. Up to four 608 caption services can be carried in a CEA-608 datastream (two per field), referred to in CEA-608 as CC1, CC2, CC3, and CC4. See 9.10 regarding consistency of CEA-708 and CEA-608 caption services and see 9.14 for information on the required caption service descriptor.

Where multiple caption services are to be encoded, the CEA-708 standard assumes that the operation will take place with a single caption encoder accepting simultaneous caption data inputs for the various services. There is no standardized method for combining DTV caption services together after encoding, although caption encoding equipment from some manufacturers allows an existing CDP to be extracted, modified with new data, and reinserted (see 6.7 for more information).

6.4 CEA-708 Caption Encoding Feeding Direct to the MPEG Encoder

The CEA-708 encoding methods described in 6.3 assume that DTVCC will be carried in the VANC of HD-SDI or SDI video signals. However, where the CEA-708 caption encoder is co-located with an MPEG encoder, the caption encoder can in some system designs feed the caption data direct to the MPEG encoder (and its associated PSIP generator for ATSC transmission). In this case, the CEA-708 caption information is typically fed over a serial data link without using VANC embedding; this technique is described in 8.3 methods 2 and 3, and 9.3 methods 3 and 4.

6.5 CEA-708 Caption Translation within the MPEG Encoder Using CEA-608 Captions from Line 21

Another alternative to the CEA-708 caption encoding in VANC described in 6.3 is possible for SD inputs to an MPEG encoder. Some MPEG encoders can accept analog NTSC or SDI video inputs with CEA-608 captions present on line 21 and internally generate the CEA-708 DTVCC captions for transmission, translated from the CEA-608 captions. See 9.3 method 5. This arrangement is expected to be an interim solution until native CEA-708 caption authoring is generally adopted, but could persist indefinitely for legacy program material already captioned in accordance with CEA-608.

6.6 Required Level of Implementation

In accordance with CEA-708, ATSC A/65, ANSI/SCTE 43 and ANSI/SCTE 54, DTV signals for broadcast and cable distribution, both for HD and SD program material with closed captions, must carry CEA-708 DTV caption data, associated CEA-608 compatibility bytes, and the caption service descriptor[2].

To enable MPEG encoders and PSIP generators to generate signals that meet the above requirements, CEA-708 caption encoders (both with and without VANC embedders) must produce the CEA-708 DTV caption data, the CEA-608 compatibility bytes, and also caption service information unless the caption service descriptor is generated and inserted later in the distribution chain. CEA-708 caption authoring systems must produce both CEA-708 and CEA-608 caption data types, and can also provide appropriate data for the caption service information.

CDPs for transporting CEA-708 caption data carry the CEA-708 DTV caption data, the associated CEA-608 compatibility bytes, and the caption service information. In accordance with SMPTE 334-2, while CEA-708 caption data, CEA-608 data, and caption service information are not all required to be present in any given CDP, it is expected that over the length of a program or commercial, all constructs typically will be present at least once. In some system architectures, however, CDPs can carry caption data without caption service information or caption service information with caption data. Time code carriage in the CDP is optional and in transports where CDPs are bound to video, time code will not necessarily be present at all.

See 9.14 for further guidance on caption service information.

6.7 Adding Additional DTV Captions after Initial Caption Authoring

Program providers are increasingly being asked to add captions in additional languages after the initial caption authoring process has been completed. Doing so requires equipment capable of extracting the original CDPs from VANC, analyzing them, adding the new language captions, and then rewriting new CDPs to VANC. Where such captioning is carried out live, the process will introduce a few frames of delay for the new captions. If this is an off-line process, the system design uses time code values to ensure that the new captions and CDPs are generated and inserted into video at the proper time to preserve the relative timing between captions and video. Care is needed to ensure that the timing and content of the CEA-608 compatibility bytes is also maintained.

Specific Information about adding additional DTVCC services to existing caption data is available from the manufacturers. Users are reminded that the caption service information (see 9.14) must correctly describe the caption service types and languages that are present.

7 Recording, Processing and Routing

This section relates to recording, processing and routing of video signals, containing closed captioning data, within a broadcast facility. This includes production and post-production facilities, network centers, and local broadcast stations.

NTSC CEA-608 captions can generally be transported over all parts of a facility and distribution chain as an integral part of an NTSC or SDI video signal. The main consideration is that video line 21 needs to be preserved through ingest, storage, and playout. With some equipment this is implicit in the design, while with other equipment (e.g., some types of video servers) special provision is made in the equipment to preserve line 21 data.

[2] The requirements for captioning television programming are covered by government regulations. Users are advised to consult the appropriate regulatory authority rules for specific information on the obligation to carry the caption service descriptor, CEA-708 captions, CEA-608 captions compatibility bytes, or to translate CEA-608 inputs into CEA-708 caption data.

Similar principles apply to transport of DTVCC carried in SDI and HD-SDI video signals. Because the SMPTE 334-1 VANC caption packets are part of the HD-SDI or SDI signal along with the video and any embedded audio, it is possible for them to automatically follow the video through most standard routing and switching and some types of processing and recording equipment. However, several issues must be considered as discussed below.

7.1 Recording

For recording systems, the integrity of the DTVCC data can be ensured by selecting a tape format and equipment or video server technology that preserves the SMPTE 334-1 data in VANC. This function can be integrated into the recording system; in other cases, as described below, it could require the use of external bridges.

It is noted that no recording equipment currently available is capable of recording and passing all VANC data from input to output. Care is needed to ensure that the relevant data packets needed to implement DTVCC in SMPTE 334-1 are preserved by the selected equipment. Individual data services are not assigned to any specific data lines in the SMPTE 334-1 standard, but it is important that the location in VANC selected for placement of the captioning data takes account of the capabilities of the recording and video processing equipment in the signal path.

For some applications, the method specified in SMPTE 337M for carriage of DTVCC in an AES3 data channel could be appropriate as an alternative to the use of VANC.

7.2 VANC Data Bridge

If the video stream is processed by equipment that changes the signal format, it will probably be necessary to bridge the captions around the conversion or compression equipment to the new video stream. One example, as shown in Figure 3, is a high definition recording system based on a standard definition SDI input recorder or video server where the compression engine function is separate from the recording function (e.g., the original implementation of HD D5 recording).

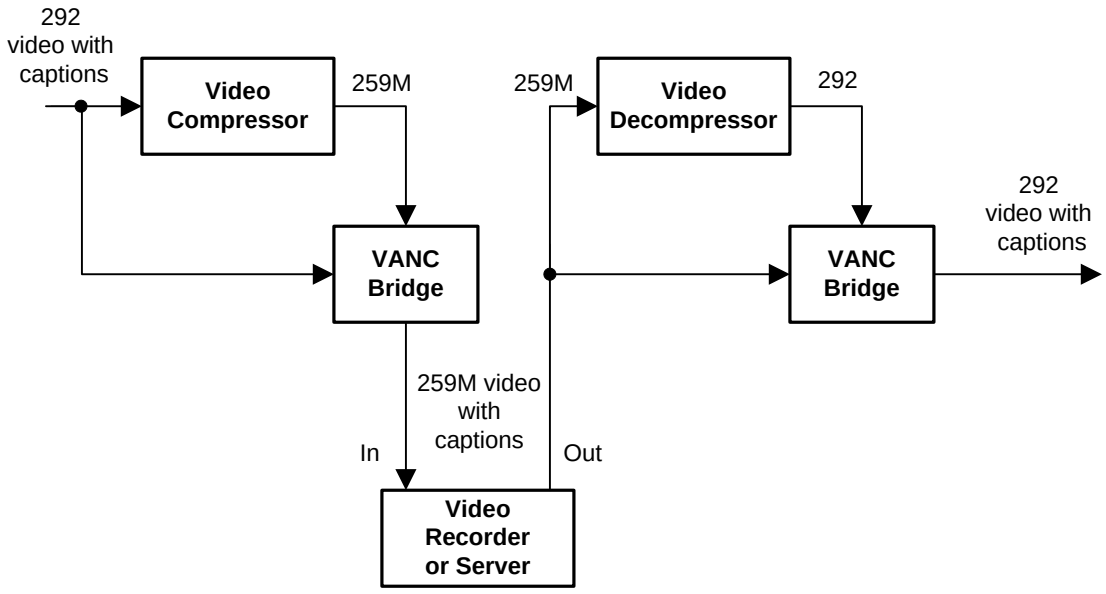

Figure 3 – Bridging VANC data around compression equipment

Because the SMPTE 291M ancillary data format used in SMPTE 334-1 VANC data mapping is identical for SMPTE 259M and SMPTE 292, using a VANC bridge unit it is possible to extract the VANC caption packets from the HD-SDI compressor input (as shown in Figure 3) and re-insert them into the SDI recorder input. Similarly, on playback, the captions can be bridged from the SDI recorder output to the HD-SDI decompressor output. The bridging equipment is required to preserve the relative timing between captions and video.

The configuration shown in Figure 3 can also apply to systems using SMPTE 305M serial data transport interface or other formats for recording or transport, rather than SMPTE 259M. Specific Information for different configurations is available from the manufacturers.

7.3 Video Processing

Various video processing equipments in the broadcast plant can delay, repeat or drop video frames and/or delete or disrupt the VANC part of the video signal. This could result in disruption to closed captions. System designers must ensure that equipment does not disrupt the flow of VANC data, or make special provision for bridging around it (see example in 7.2 where the VANC data would be completely lost in the absence of a bridge). The bridging equipment must preserve the relative timing between captions and video as far as possible.

7.4 Switching Considerations

When switching between two sources with CEA-708 or CEA-608 captions, there can be some effect of the old caption context on the new captions until the context is retransmitted. Thus, the new captions could be displayed without their native attributes until the new caption context is sent to the receiver. This situation is similar to what happens when a receiver first tunes to a station, i.e., it must acquire the caption context before it can display the captions correctly.

To minimize this effect, it is recommended that caption-originating production facilities:

a. Keep the caption context static except when there is a need for a change related to the program content (e.g., changing font colors for different talent).

b. Re-establish the context frequently. The recommended practice for caption creation and encoding is to include the context with each caption.

Typically, these practices are implemented in caption authoring procedures.

8 Contribution and Distribution

Contribution, in this context, generally refers to program feeds into network facilities, or from a program source direct to a local broadcast station. Distribution refers to program feeds from a network center to local broadcast stations. The considerations in this section apply equally to contribution and distribution feeds if they carry closed captioning. They also apply equally to distributors targeting either terrestrial emission, cable, or satellite distribution systems.

The main consideration for NTSC CEA-608 captions is that video line 21 needs to be preserved through the transport mechanism. Typically, distribution systems used for broadcast purposes meet this requirement and are transparent to line 21 data.

The same principles apply to transport of DTVCC that are carried within the plant in the VANC of HD-SDI and SDI video signals. When DTVCC are transported on systems outside the plant, whether based on tape, computer files, copper cable, fiber, microwave or satellite, several issues must be considered as discussed below.

8.1 Tape Interchange

In cases where the interchange medium for program material is videotape, as noted in 7.1 the integrity of the DTVCC data must be ensured by selecting a tape format and equipment that preserve the SMPTE 334-1 data in VANC. This applies to recorders for both HD and SD program material carrying DTVCC.

8.2 File-Based Delivery

In addition to other program delivery methods, programs can be delivered by electronic means as computer files. The MXF and GXF file formats standardized by SMPTE specify the carriage of closed captions, both CEA-608 and CEA-708, within the file structure. See SMPTE 377M for MXF and SMPTE 360 and SMPTE RDD 14 for GXF. Other file formats that are not standardized might or might not carry closed captions.

Files that use MPEG-2 video compression encoding may carry closed caption data in the video user data, as per ATSC A/53 Part 4. Similarly, files that use AVC compression might carry closed caption data in the supplemental enhancement information (SEI), as per ATSC A/72 Part 1. See also 8.3 below.

8.3 Contribution/Distribution Compression Systems

Note: The techniques described in this section for “MPEG compression systems” apply equally to MPEG-2, AVC, or SMPTE VC-1 video codecs and, in principle, to other compression codecs.

If the program is distributed through a medium that involves MPEG compression VANC data is not carried directly with the compressed video, therefore the caption data must be bridged from the VANC packets of the input HD-SDI or SDI video signal to a data channel in the MPEG multiplex. In general, there are three different (and potentially co-existing) ways to do this, as follows:

1. At the uplink encoder the VANC data is packaged into the transport stream multiplex as a private data service on dedicated PIDs (packet ID), including MPEG presentation time stamps (PTS). At the receiving end, the professional decoder (usually an integrated receiver decoder, IRD) then extracts the CDP packet stream from the multiplex and re-inserts it as VANC SMPTE 334-1 data into the HD-SDI or SDI video signal resulting from MPEG decoding. Frame accuracy of the VANC data to the video must be preserved through the PTS. The bitstream syntax and semantics for two implementations of such a system are documented in SMPTE 2038 and RDD 11.

2. Alternatively, for captions generated direct from a CEA-708 caption encoder, an EIA/TIA-232 serial representation of the captions in a non flow-controlled form such as the “Grand Alliance” protocol (see Annex A of SMPTE RP 2007, Closed-Caption CDP and "Grand Alliance" Serial Interfaces for DTV) could be packaged into a private data service and reconstructed and output as EIA/TIA-232 data at the decoder or IRD.

3. The uplink encoder could also take the caption data (whether it comes in VANC with a video signal or as EIA/TIA-232 serial data direct from a CEA-708 caption encoder) and pack it into the video user data according to CEA-708 and ATSC A/53. At the receiving end, the IRD might then output a transport stream including the caption data, or it might also extract the caption data and reformat it into VANC SMPTE 334-1 data inserted in the HD-SDI or SDI video signal resulting from MPEG decoding, or output it on a serial link.

Method 1 is shown in Figure 4. The separate SMPTE 334-1 disembedder and embedder units shown with dotted lines will be required if the MPEG encoder/multiplexer does not have the capability to extract the VANC with caption CDPs, and/or the IRD does not have the capability to insert the CDP into the VANC of the video output.

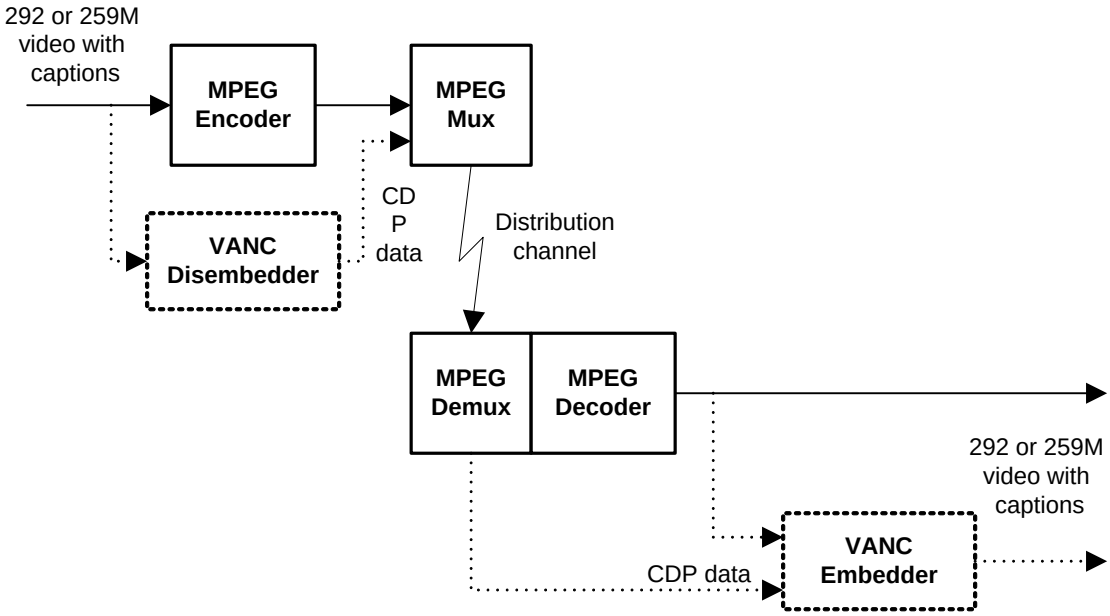

Figure 4 – Compressed contribution or distribution

Points to note are:

a. All three of the above methods can be implemented at the uplink system, and the most useful one(s) selected for use at the downlink site. All will equally support CEA-708 captioning data and the CEA-608 compatibility bytes at the same time. Only method 1 supports CDPs, which can include caption service information and time code. If methods 2 or 3 are used, other arrangements must be implemented for feeding caption service information to the PSIP generator and emission encoder (see 6.6 and 9.14). Also, with methods 2 or 3 the system design cannot rely on timecode being passed with the captions through the chain.

b. Methods 1, 2, and 3 are supported with MPEG transmission at bit rates including, and higher than, the 19.39 Mbps ATSC emission standard, such as the 45 Mbps rate typically used for network distribution using MPEG-2. For methods 1 and 2, details such as PID allocations will be specific to individual manufacturers and the decoder will need to be configured to work with the uplink encoder.

c. Method 3 is similar to the emission station captioning and ATSC encoding arrangement shown in Figure 5, and is likely to ensure interoperability between the encode and decode ends of the distribution link. It is usually used for network distribution at 19.39 Mbps if at the receiving station end the bitstream has to be capable of being passed straight through to the DTV transmitter. It can also be used when the bitrate differs from the ATSC nominal rate, using the video outputs with captions data in VANC or feeding caption data via serial link to an MPEG emission encoder.

In each case the equipment bridging the caption data through the MPEG process must preserve the relative timing between captions and video.

9 Emission

Figure 5 shows a possible configuration of captioning-related equipment in a local station with ATSC/DTV transmission, and an (optional) NTSC output that might provide a direct feed to a cable headend or, in some cases, could feed an analog NTSC transmitter. This is intended only to illustrate the captioning arrangements discussed below and is not intended as a full system diagram.

The drawing shows a DTV system for HD only. A similar arrangement could be used for SD DTV using SMPTE 259M in place of SMPTE 292. In that case up and down conversion to and from the NTSC feeds is not required.

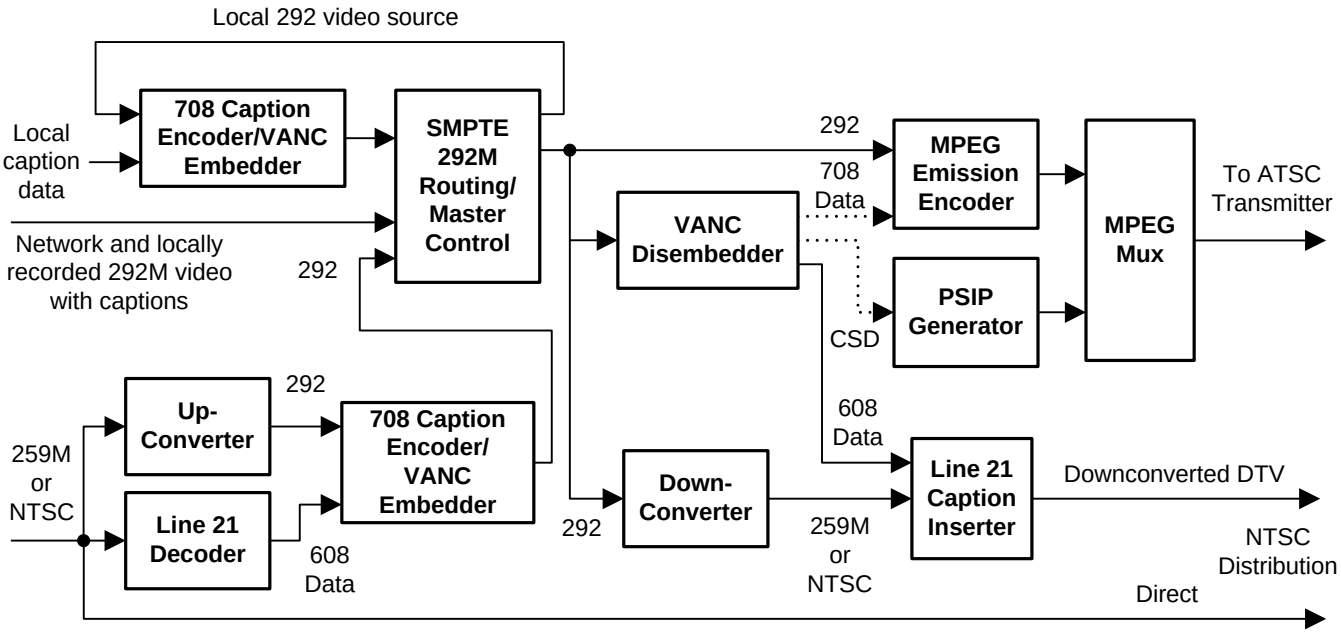

Figure 5 – Emission station captioning equipment

9.1 NTSC Output

For an NTSC distribution output (if present), there are two main arrangements. These typically comprise:

1. A legacy NTSC station master control with legacy network distribution or other external program source and local sources, or

2. A standard definition (possibly down-converted) feed from one of the DTV program services.

For the first case, NTSC caption arrangements will be unchanged from the legacy arrangement. For the second case, line 21 caption data will usually need to be inserted, see 9.8.

9.2 ATSC Station Output

For the ATSC station output there are three main arrangements, these typically comprise:

1. A dedicated digital (possibly HD) station master control or simple switcher with a (digital network distribution and local live and recorded sources. The DTVCC are usually already present on the live network and recorded sources and are added to the live local sources;

2. A feed (possibly upconverted) from the NTSC network distribution and master control. In this case the line 21 captions will need to be translated/transcoded to CEA-708;

3. A bitstream received from the network already encoded in ATSC transmission format (at 19.39 Mbps) and ready to send to the DTV transmitter (with new PSIP information). In this case, DTVCC must already be present in the bitstream. This arrangement is not shown in Figure 5.

Some stations will have combinations of these arrangements, depending on network DTV distribution arrangements and the number of DTV program services, and they could change dynamically during the day.

9.3 Captioning Inputs to an MPEG Emission Encoder

The task of embedding the caption data into the ATSC stream in accordance with A/53 Part 4 and CEA-708 falls to the MPEG emission encoder. Several different methods can be used for getting the caption data into the encoder, depending on system configuration and manufacturer and the source of the caption data, as follows:

1. Most emission encoders have the capability of extracting CDPs from SMPTE 334-1 VANC packets; in that case, the "708 Data" connection shown dotted in Figure 5 is not required.

2. Some emission encoders require a separate VANC disembedder for SMPTE 334-1 data to extract the CDP stream from VANC in an HD-SDI or SDI video signal and transform it to the data format expected by the MPEG emission encoder, as shown in Figure 5. Most ATSC MPEG emission encoders have a caption input port (usually TIA-574 or TIA-232 asynchronous serial). This port accepts data in one of two documented formats: SMPTE 333, which is known as the "pull" protocol; or the so-called "Grand Alliance" format, also known as the "push" protocol, which is described in SMPTE RP 2007, "Closed-Caption CDP and 'Grand Alliance' Serial Interfaces for DTV".

3. Alternatively, as shown in Figure 6, a local CEA-708 caption encoder without VANC embedding can be used to feed the MPEG encoder using the serial feed as described in method 2 above and taking caption data from a captioning computer or possibly from a live authoring workstation. Figure 6 does not show the video routing and possible special arrangements that could be needed for switching captioning data signals for different sources.

4. Alternatively, also as shown in Figure 6, a CEA-608 to CEA-708 translator/transcoder unit without VANC embedding can be used to feed the MPEG encoder with captions originating from a legacy line 21 caption source, using the serial feed as described in method 2 above.

5. Alternatively, as described in 6.5, some MPEG encoders can accept analog NTSC or SDI video and generate the CEA-708 DTVCC captions for transmission from CEA-608 legacy captions present on line 21 of the incoming feed, without using an external CEA-708 caption encoder.

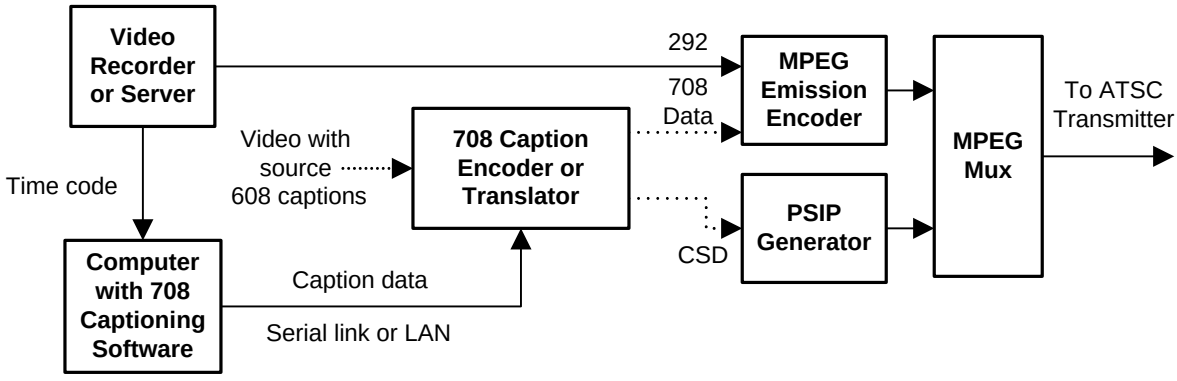

Figure 6 – Alternative captioning feed to emission encoder

Points to note:

a. Methods 1 and 2 allow all captioned programs to be routed and processed in the same way, regardless of their origin. They require that VANC-embedded DTVCC are present on the HD-SDI or SDI inputs and output from the DTV master control switcher. This eliminates the need for special switching for caption data signals. Arrangements for embedding DTVCC in local live sources and NTSC upconverted sources before reaching the master control switcher are discussed below.

b. Methods 3 and 4 allow local captioning of program material without VANC-embedded DTVCC in the case where captions are generated live or are bridged from another source (e.g., a simulcast NTSC feed of the same program).

c. Method 5 is typically used where stations are encoding an SD DTV program from an analog or SDI program originally intended for NTSC transmission.

d. A combination of the above methods can be used. This requires that special switching (not shown in the drawings) is used to route to the MPEG encoder the appropriate caption data for DTV network, local sources, or NTSC converted sources.

e. SMPTE 333 refers to the device feeding the caption data to the video (emission) encoder as a “caption server”. In that usage, the term applies to all the devices feeding the MPEG encoder as mentioned in methods 2, 3, and 4 above and differs from the use of that term in sections 6 and 9.13 of this document.

f. One caption equipment manufacturer has implemented a non-standard “push-mode” extension of the SMPTE 333 protocol for private transport of CEA-708 data where SMPTE 333 packets are sent without the requirement for the receiving device to send requests (i.e. SYN bytes). As indicated in Note 1 to Annex A of SMPTE RP 2007, there is also a non-standard implementation of the Grand Alliance protocol by one manufacturer. Users are advised to check with the manufacturers of the caption encoder and MPEG encoder to ensure that the format used by each is compatible.

g. The length of VANC CDPs can vary for differing production formats and with film-based materials. The MPEG encoder must handle correctly each of the packet lengths extracted, even if they are in error or do not correspond to the frame rate to be encoded.

9.4 Captioning in AVC Bitstreams

Discussions of the specifics of AVC captioning for ATSC systems and cable distribution are found in 4.3.3 and 4.4.6. System arrangements for feeding program and captions to the AVC encoder are typically the same as those described in 9.2 and 9.3.

9.5 Captioning for ATSC Mobile DTV

Discussions of the specifics of ATSC Mobile DTV captioning are found in 4.3.4. Program material for Mobile/ DTV usually originates as HD or SD format video and arrangements for feeding program and captions to the Mobile DTV AVC encoder are typically the same as those described in 9.2 and 9.3.

9.6 Local Live DTV Captions

Live programming can be captioned using a CEA-708 caption encoder with SMPTE 334-1 VANC embedding, as discussed in 6.3, but with live video and with caption intentions generated in real-time. This is shown at the upper-left side of Figure 5. Alternatively, the CEA-708 caption encoder can feed the MPEG emission encoder directly as in 9.3 method 3.

If the local programming originates as an NTSC source, existing legacy arrangements for live CEA-608 captioning arrangements can be used, with bridging to DTVCC as in 9.7. All bridging equipment must preserve the relative timing between captions and video.

9.7 NTSC – DTV (Up-Conversion)

Some HD-DTV programs are produced by up-conversion from an analog NTSC or SDI program containing CEA-608 captions. In this case, the captions must be bridged from the analog NTSC or SDI line 21 to CDPs in the VANC of the HD-SDI stream. This can be accomplished by having a decoder extract the CEA-608 data from line 21 and feed it to a CEA-708 caption encoder/VANC embedder, as shown at the lower-left side of Figure 5.

Two message formats adopted by different caption equipment manufacturers for this transfer are described in Annex C.

Similar techniques can be used to insert DTVCC in the VANC of SD programs for DTV transmission, without up-conversion.

9.8 DTV – NTSC (Down-Conversion)

It is sometimes necessary to derive an NTSC video signal from an HD-SDI stream. The CDPs carried in SMPTE 334-1 VANC packets in the HD-SDI stream contain CEA-608 data, which can be embedded into line 21 of the analog NTSC or SDI video. This can be achieved by having a 334-1 VANC disembedder extract the data from the HD-SDI stream and feed the CEA-608 caption data via a serial port to an analog NTSC or SDI line 21 closed caption inserter, as shown in the lower-right portion of Figure 5.

Two data formats adopted by different caption equipment manufacturers for this transfer are described in Annex C.

Similar techniques can be used to derive an NTSC video signal from an SDI stream with CDPs in VANC, without down-conversion.

9.9 Completeness of CEA-608 Captions, Text, and Program Data

Whichever method is used for feeding caption data to the MPEG emission encoder, it is important that the CEA-608 compatibility bytes data provided for the CEA-708 DTVCC data construct include all the data that would normally have been carried in the broadcaster's previous NTSC service, which could include all four caption services, all four text services and all XDS bytes. This ensures that any NTSC signal that is recreated from the ATSC broadcast signal will include line 21 populated with all appropriate captioning and other data.

Section 7 of the ATSC A/79, Recommended Practice for Conversion of ATSC Signals for Distribution to NTSC Viewers provides guidance to broadcasters on handling CEA-608 data. It notes that where such data is generated and inserted into the CDPs depends on the operational practices of the broadcaster. If an NTSC signal chain including CEA-608 line 21 data is maintained within the broadcast facility, the CEA-608 data can be bridged to the CEA-708 caption encoding device that inserts the CDPs into the SDI or HD-SDI video signal. Where the line 21 infrastructure is no longer present, alternative paths must be established to provide the necessary 608 data to the CEA-708 caption encoder.

A/79 also notes that, as with previous NTSC program distribution systems, some CEA-608 data can originate at the network release facility and some or all at the local broadcast station. Current generation CEA-708 caption equipment allows for CDPs distributed with a video signal from a network to be modified if necessary downstream at the local station so they include valid and current CEA-608 data for the program being transmitted, including XDS data such as content advisory (V-Chip), network name, time-of-day, and other "air time specific" XDS packets.

9.10 Consistency of CEA-708 and CEA-608 Caption Services

The CEA-708 and CEA-608 captions are expected to be consistent with each other and both relate to the same program with which they are carried. In accordance with CEA-708 and CEA-608, DTV Service #1 and CEA-608 caption service CC1 are the services designated as the Primary Caption Service. These both contain the verbatim or near-verbatim captions for the primary language being spoken in the accompanying program audio and need to be in sync with the sound.

DTV Service #2 and CEA-608 caption CC3 are designated as the Secondary Language Service[3]. These both contain captions in a secondary language, which are either translations of the captions in the Primary Caption Service or captions of a Secondary audio service, and also need to be in sync with the sound, although they could be in a different language.

The other DTV and CEA-608 caption service sub-channels are not pre-assigned.

9.11 Consistency of XDS Data and PSIP Data

Broadcasters must ensure consistency between the content advisory bytes ("V-chip") when carried in CEA-608 XDS data and the content advisory information carried in the EIT for the same broadcast event. Depending on the operational practices of the broadcaster, it could be desirable to provide equipment that reads EIT data from the PSIP for the broadcast and verifies or inserts appropriate data into incorrect or missing XDS bytes.

9.12 ANSI/SCTE 127 Data

ATSC MPEG emission encoders can also encode other VBI and line 21 payloads as defined in ANSI/SCTE 127 and as specified in ATSC A/99. The source, carriage and disposition of such data are covered in the named standards and it is not part of the CEA-608 data as described in this guideline. Some legacy implementations of AMOL, used for audience measurement purposes, possibly will still be carried in line 21 by some operators.

9.13 Server-based Captioning

The techniques described above for carrying captioning data in VANC of the video signal ("sticky data") are appropriate for many applications of DTV caption distribution. An alternative server-based captioning architecture, where the caption data is distributed separately from video ("slippery data") could have advantages for some station operations, and is particularly applicable in facilities using video-servers for program origination. A server-based system avoids issues associated with recording and distributing VANC caption data, and can provide greater flexibility for some applications. It allows programs to be edited after captioning has been carried out while maintaining captioning integrity, provides enhanced capability for captioning multiple versions of programs; e.g., for different distribution platforms, and for recording live captioning with later editing and re-use.

In this type of system, instead of inserting caption data into the video at some upstream point for distribution with the video program, caption files are generated and then distributed separately from the video, typically over a computer network. A separate caption file for each program is stored on a captioning computer (server) at the emission station that runs software to manage the various caption files and playout. Software running on the captioning server, or on other networked computers, allows a program playlist to be established. As programs are played out, time code from the video server triggers captions from the caption server to be fed to the caption encoder and inserted into the transmission video feed. Other architectures are possible, including systems where the captions are added to the video media on the video server, without using a separate caption encoder. Server-based, time-of-air caption encoding techniques are largely proprietary, although often utilizing some or all of the systems mentioned in this guideline. Information on specific features and arrangements is available from the manufacturers.

[3] Some broadcasters prefer CC2 for secondary language captioning in CEA-608 because some consumer decoders do not support CC3. The downside is that CC2 is not guaranteed to be in sync with the audio.

Server-based captioning is applicable to both CEA-608 captioning for NTSC and CEA-708 captioning for DTV. Such systems use basically the same interfaces between the captioning computers and caption encoders, and between caption encoders and MPEG encoders, as described elsewhere in this guideline.

9.14 Caption Service Descriptor

Up to 16 services can be announced in the caption service descriptor (CSD), which identifies each service as "digital" 708 or "line-21" 608 type captions. For 708 services only, the CSD identifies each service with a caption service number in the range of 1 – 63 and must identify the dominant language for each service; it can also indicate whether the captions are wide aspect ratio and/or easy reader. For 608 services, there is no mechanism for signaling of caption service number, language, or other information. There is a "line21_field" flag defined for the caption service descriptor but A/65 states that its use is deprecated and it may be set or clear.

The CSD carrying caption service information is required to be carried in the DTV bitstream as set out in ATSC A/65, ATSC A/53, A/153, CEA-708, and ANSI/SCTE 54. For systems using CDPs to carry caption service information, the caption service information can be extracted from the CDP when the VANC data is disembedded and then fed to the PSIP generator and, where appropriate, the MPEG emission encoder as shown in Figure 5. That equipment is required to enter the CSD into the current program EIT-0 (or Service Map Table, SMT, in the ATSC Mobile DTV standard) and (optionally) the PMT table entries, as required by the standards mentioned above. Local CEA-708 caption encoders feeding the MPEG emission encoder and PSIP generator directly without using VANC, as shown in Figure 6, also need to generate the caption service information.

9.14.1 Caption service information in CDPs

As specified in SMPTE 334-2, caption service information for one or more services (up to 16) can be spread over a sequence of CDPs. System designers and users are alerted to the fact that the data for the CSD loop for different caption services has to be parsed from the appropriate CDP or CDPs for each service.

The expectation from CEA-708 and ATSC A/65 is that the caption service information in broadcast signals will be constant over some well-understood time period for the application, such as a program event duration, and typically will not change for a short duration event such as a commercial. The caption service information needs to be sent to the PSIP generator far enough in advance of an emission of an EIT-0 that it can be included in the EIT-0 (or SMT for ATSC Mobile DTV) and the PMT as required. SMPTE 334-1 does not specify a repetition rate for caption service information sent in the CDPs, but it is suggested that it be sent at least twice per second. This would enable downstream VANC-aware equipment, even if there is a discontinuity in the video signal distribution, to extract current caption service information from the CDPs and send it to other equipment that inserts the contents with the descriptor header to populate CSD fields in the ATSC bitstream.

9.14.2 Alternative arrangements for caption service information

It is noted that methods for originating caption service information are not standardized, and different users and manufacturers can have different arrangements for generating and distributing caption service information. It cannot be assumed that all caption encoders generate complete and correct caption service information for the caption services as required for insertion in the PSIP tables.

How and where the caption service information data originates will depend on the system architecture adopted and the operational practices of the broadcaster. Many program management and traffic database systems used by broadcast networks and stations carry caption service information for current and future programs and segments. For some system architectures this is the preferred source of caption information for the PSIP generator, rather than relying directly on caption service information carried in the CDP. Information for caption service descriptors to be entered in PSIP tables for future events has, of necessity, to be generated and distributed separately from the captioned video service. Nevertheless, it is recommended that,

where possible, captioned program material using CDPs carry complete and accurate caption service information, bound with the video in the CDPs.

Users are advised to check with the manufacturers of the caption encoder, PSIP generator and MPEG encoder, and also of program management and/or traffic systems, to determine arrangements needed to meet the requirements of the standards

Points to note:

a. Annex B describes a message format that was developed by one manufacturer to allow the CSD and also the content advisory descriptor to be provided from a caption encoder or VANC disembedder to a PSIP generator upon request. This method for communicating caption service information has not, however, been implemented in deployed systems.

b. ATSC A/76 is an XML-based open standard for communication of PSIP/PSI-related metadata between equipment in the facility. This standard facilitates transfer of both current and future program metadata (including the caption service descriptor) to the PSIP generator, and is the preferred method of communicating PSIP metadata when not bound to video or audio.

c. The caption service information defined in ATSC A/76 is also included in the SMPTE 2021 Broadcast Exchange Format (BXF) standard, which is an alternative preferred method of communicating caption service information when not bound to video or audio.

10 Monitoring

Monitoring of CEA-608 and 708 CEA-708 caption encoding and content must be carried out at appropriate points in the distribution chain to check for:

- Presence of captioning information
- How captions will be displayed on a consumer receiver (basic caption functionality)
- Video/caption synchronization
- Regulatory compliance

In particular, it is strongly recommended that caption data is checked when the captions are generated and when programs are delivered. Master control facilities need monitoring for closed captioning data and functions, both for incoming program material and for transmitted broadcast signals.

10.1 608 Caption Monitoring

608 caption monitoring requires a professional line 21 caption decoder, with analog, SDI, or RF inputs, with the different inputs being needed for different points in the chain before and after transmission. Such units provide comprehensive monitoring and reporting functions for captioning and other associated data functions and can incorporate on-screen visual display of captioning information and/or alarm, logging and remote reporting functions for non-standard and fault conditions.

A standard NTSC television receiver/monitor with caption decoding capability allows consumer-level captioning and associated functions to be monitored where the analog NTSC video or off-air signal is available, and provides an inexpensive way of constantly monitoring off-air basic performance.

10.2 708 Caption Monitoring

708 caption monitoring has multiple requirements:

1. 708 caption authoring systems need to provide the ability to view both CEA-708 and the associated CEA-608 caption streams simultaneously as they are produced to ensure that both types of captions are useable.

2. Arrangements are needed to monitor full CEA-708 captioning functions prior to transmission. Some CEA-708 encoders themselves provide the capability for monitoring and verifying caption data embedded in the video signal. When monitoring downstream in the VANC caption distribution chain, monitoring requires use of a VANC disembedder with appropriate caption decoders to examine and view the CEA-708 and CEA-608 captioning and associated data functions. Such CEA-708 DTVCC monitoring equipment is now available and provides comprehensive monitoring and reporting functions for DTV captioning; functions can incorporate on-screen visual display of captioning information and/or alarm, logging and remote reporting functions for non-standard and fault conditions. Ideally, monitoring equipment will provides three levels of DTVCC checking:

    - Visual display of the rendered captions,
    - Logging and consistency checking of the decoded stream commands, and
    - Logging and error checking of the actual raw stream data bytes.

    With these levels of checking, the integrity of the caption data can be followed through the entire decoder pipeline to determine the location of any fault or misinterpretation of the data.

3. It is necessary to verify correct insertion of the CEA-708 and CEA-608 caption data in the outgoing DTV bitstream. Presence of the caption data and the caption service descriptor in the MPEG transport stream can be checked using an MPEG stream monitor. The correct synchronization of captions with the associated video frames must also be verified.

4. It is essential to monitor the broadcast CEA-708 captions using an off-air DTV receiver with CEA-708 closed captioning decoding capability. It is important to view the captions on a typical consumer television set, as used by the intended audience. However, professional receivers are also useful as they can provide more comprehensive functionality (and reporting capabilities) than typical consumer receivers. Some professional receivers, in addition to an RF input, have a transport stream input that can be used to monitor the bitstream output from an ATSC (MPEG) encoder prior to sending to the transmitter. This capability will assist in system trouble-shooting. Some professional receivers also provide a demodulated form of the bitstream, typically on a DVB-ASI (Digital Video Broadcasting Asynchronous Serial Interface) output) which can then feed an MPEG stream analyzer.

5. Monitoring of cable television captioning requires use of a suitable cable television interface unit or integrated receiver able to select and display the DTV captions.

6. DTV caption monitoring must include the associated CEA-608 caption data using the derived NTSC signal generated by the DTV receiver to feed CEA-608 monitoring equipment as described in 10.1.

Annex A Bibliography (Informative)

At the time of publication, the editions indicated in this bibliography were valid. All standards are subject to revision, and users of this guideline are encouraged to investigate and use the most recent edition of the standards indicated below.

AES3-2003, Serial Transmission Format for Two-Channel Linearly Represented Digital Audio Data

ANSI/SCTE 20-2004, Methods for Carriage of Closed Captions and Non-Real Time Sampled Video

ANSI/SCTE 21-2001r2006, Standard for Carriage of NTSC VBI Data in Cable Digital Transport Streams

ANSI/SCTE 40-2004, Digital Cable Network Interface Standard

ANSI/SCTE 43-2005, Digital Video Systems Characteristics Standard for Cable Television

ANSI/SCTE 54-2007, Digital Video Service Multiplex and Transport System for Cable Television

ANSI/SCTE 127 2007 Carriage of Vertical Blanking Interval (VBI) Data in North American Digital Television Bitstreams

ANSI/SCTE 128-2008, AVC Video Systems and Transport Constraints for Cable Television

ANSI/SCTE 157-2008, VC-1 Video Systems and Transport Constraints for Cable Television

ATSC A/53 (2007/2009), ATSC Digital Television Standard, Parts 1 - 6

ATSC A/65C (2006), Program and System Information Protocol for Terrestrial Broadcast and Cable, Rev. C, with Amendment No. 1

ATSC A/72 (2008), Part 1: Video System Characteristics of AVC in the ATSC Digital Television System

ATSC A/72 (2008), Part 2: AVC Video Transport Subsystem Characteristics

ATSC A/76B (2008), Programming Metadata Communication Protocol Standard

ATSC A/79 (2008), Recommended Practice for Conversion of ATSC Signals for Distribution to NTSC Viewers

ATSC A/99 (2008), Carriage of Legacy TV Data Services

ATSC A/153 (2009), ATSC Mobile DTV Standard, Parts 1 - 8

CEA-608-E (2008), Line 21 Data Services

CEA-708-D (2008), Digital Television (DTV) Closed Captioning

CENELEC EN 50083-9-2002, Cable Networks for Television Signals, Sound Signals and Interactive Services; Part 9: Interfaces for CATV/SMATV Head-ends and Similar Professional Equipment for DVB/MPEG-2 Transport Streams

DVB/ETSI TS 101 154 v 1.8.1 (2007), Specification for the use of Video and Audio Coding in Broadcasting Applications based on the MPEG-2 Transport Stream

FCC 00-259-2000, Report and Order, Closed Captioning Requirements for Digital Television Receivers, Federal Communications Commission

FCC 05-142-2005, Notice of Proposed Rulemaking, Closed Captioning of Video Programming, Federal Communications Commission

FCC 08-255-2008, Declaratory Ruling, Order, and Notice of Proposed Rulemaking, Closed Captioning of Video Programming; Closed Captioning Requirements for Digital Television Receivers, Federal Communications Commission

SMPTE 12M-1-2008 Television – Time and Control Code

SMPTE 170M-2004, Television – Composite Analog Video Signal – NTSC for Studio Applications

SMPTE 259M-2008, Television – SDTV Digital Signal/Data – Serial Digital Interface

SMPTE 291M-2006, Television – Ancillary Data Packet and Space Formatting

SMPTE 292-2008, 1.5 Gb/s Signal / Data Serial Interface

SMPTE 305M-2005, Television – Serial Data Transport Interface (SDTI)

SMPTE 333-2008, DTV Closed-Caption Server to Encoder Interface

SMPTE 334-1-2007, Vertical Ancillary Data Mapping of Caption Data and Other Related Data

SMPTE 334-2-2007, Caption Distribution Packet (CDP) Definition

SMPTE 360-2009, General Exchange Format

SMPTE 377-1 2009, Material Exchange Format (MXF) – File Format Specification

SMPTE 2021-2008, Broadcast Exchange Format (BXF)

SMPTE 2038-2008, Carriage of Ancillary Data Packets in an MPEG-2 Transport Stream

SMPTE 2046-1-2009, Specifications for Safe Action and Safe Title Areas for Television

SMPTE RDD 11-2007, Bitstream Syntax and Semantics for Carriage of HDSDI Ancillary Data in an MPEG-2 Transport Stream

SMPTE RDD 14-2007, General Exchange Format-2

SMPTE RP 202-2008, Video Alignment for Compression Coding

SMPTE RP 207-2005, Transport of Program Description Data in Vertical Ancillary Data Packets

SMPTE RP 218-2002, Specifications for Safe Action and Safe Title Areas for Television Systems

SMPTE RP 2007-2007, Closed-Caption CDP and "Grand Alliance" Serial Interfaces for DTV

TIA-232 (2002), Interface Between Data Terminal Equipment and Data Circuit-Terminating Equipment Employing Serial Binary Data Interchange

TIA-574 (2003), 9-Position Non-synchronous Interface Between Data Terminal Equipment and Data Circuit-Terminating Equipment Employing Serial Binary Data Interchange

Annex B Descriptor Enquiry Message Format (Informative)

The following describes two commands that can be sent to a SMPTE 334-1 (VANC) decoder (disembedder) through an Ethernet or TIA-574 connection to allow a PSIP generator to retrieve descriptors needed for the broadcast bitstream. This description is based on commands implemented by one manufacturer of a SMPTE VANC decoder.

Each command begins with a Start-of-Header character (SOH = 01 hexadecimal), followed by the actual message consisting of ASCII characters, and ends with a Carriage Return character (CR = 0D hexadecimal).

All characters in these commands and responses are transmitted with odd parity over TIA-574 connections, and no parity for Ethernet connections. The data rate is bounded by the physical interface that is used.

Caption service information enquiry

This command can be used by a PSIP generator to retrieve the Caption Service Information from the Caption Distribution Packet in the video signal.

Command format: <SOH> N_E1 <CR>

The components of this message are as follows:

<SOH>	SOH character (0x01)
N_E1	ASCII characters "N", "_", "E", and "1"
<CR>	CR character (0x0D)

Response: When this command is received, the complete set of service information most recently extracted is returned in the following format:

<LF> {ccsinfo_section} {checksum} <CR>

The components of this message are as follows:

<LF>	Line Feed (0x8A)
{ccsinfo_section}	the data formatted according to Table 6 in SMPTE 334-2
{checksum}	an 8-bit checksum that is calculated to satisfy the following: the modulo-256 sum of all the 8-bit values between the <LF> and <CR> characters, including the checksum, is equal to zero.
<CR>	Carriage Return (0x0D)

Note that every 8-bit value in {ccsinfo_section} and {checksum} is sent as 2 ASCII-encoded hexadecimal digits (with odd parity when the connection is TIA-574), with the most-significant digit sent first. For example, the first byte in Table 28, ccsvcinfo_id, which has a value of 0x73, would be represented by the 2 bytes: 0x37, 0xB3 (0x37 is the ASCII code for the digit ‘7’, and 0xB3 is the ASCII code for the digit ‘3’ with odd parity added). Also, there are no space characters between sections of the response message such as {ccsinfo_section} and {checksum}. The checksum is calculated from the 8-bit data values, before they are converted to pairs of ASCII characters.

Content advisory enquiry

This command can be used by a PSIP generator to retrieve the Content Advisory Program Descriptor packet from the video signal.

The formats of the command and response follow the same rules as the Caption Service Information Enquiry, including the computation of the checksum and the parity.

Command format: <SOH> N_E2 <CR>

The components of this message are as follows:

<SOH>	SOH character (0x01)
N_E2	ASCII characters "N", "_", "E", and "2"
<CR>	CR character (0x0D)

Response: When this command is received, the Content Advisory packet most recently extracted is returned in the following format:

<LF> {content_advisory_descriptor} {checksum} <CR>

The component {content_advisory_descriptor} is the data formatted according to table 6.25 in ATSC A/65.

The Content Advisory message is sent to the PSIP generator as ASCII-encoded hexadecimal values, similar to the Caption Service Information message.

Annex C Field-Marked Message Format for CEA-608 Data (Informative)

These message formats can be used to transfer CEA-608 data between caption decoders and caption encoders, for example in support of video up and down-conversion. The following methods are implemented by two different manufacturers of caption encoding equipment.

Method 1

The data format for this transfer consists of a three-byte sequence for each CEA-608 triplet in the CDPs. Each three-byte sequence consists of a field marker byte (0x81 and 0x82 hexadecimal, for fields 1 and 2 respectively), followed by two bytes of CEA-608 data. The values of the marker bytes are selected to have even parity, which distinguishes them from caption data bytes that always have odd parity. If a field contains no CEA-608 data, this is indicated by a three-byte sequence consisting of a field marker byte of 11 or 12 hexadecimal, for fields 1 and 2 respectively, followed by two null bytes with odd parity (i.e. 0x80 hexadecimal). Note that the condition of no CEA-608 data refers to either of the following cases:

a) For an NTSC signal, there is no valid Clock Run-in.

b) For the VANC caption data from a SMPTE 292 signal, the cc_valid bit for the pair of CEA-608 data bytes is zero, or there is no Caption Distribution Packet present.

Method 2

This format uses a serial protocol with 8 bits, no parity, and 1 stop bit (8,N,1). This is typically set to 9600 baud but other baud rates can be selected. The protocol consists of CEA-608 byte pairs. CEA-608 data requires only 7 bits, so bit 7 (MSB) of each byte is allocated as a field indicator. MSB is cleared for field 1 and set for field 2. Null byte-pairs are not transmitted. However, null bytes that occur as part of XDS data packets are transmitted.