

TOS Pt.1



PBS Technology & Operations

TECHNICAL OPERATING SPECIFICATION

Part 1: Program Submission

2020 Edition

Table of Contents

1	Scope and Purpose	1
1.1	Purpose.....	1
1.2	Requirements	1
1.3	Producer Responsibilities	1
1.4	File Based Media.....	1
2	Video	2
2.1	Video Image Quality	2
2.2	Video Definition.....	2
2.2.1	HD Image.....	2
2.2.2	SD Upconvert	2
2.3	Video Anomalies.....	2
2.3.1	Compression Artifacts.....	2
2.3.2	Analog Artifacts.....	2
2.3.3	Non-Image Content.....	2
2.4	Video Levels.....	2
2.4.1	Conversion	2
2.4.2	Color Gamut	3
2.4.3	Black Level.....	3
2.5	Video Image.....	3
2.5.1	Image Format	3
2.5.2	Pillarbox Content.....	3
2.5.3	4:3 Protect.....	3
2.5.4	Safe Areas.....	3
2.5.5	Subtitles.....	4
2.5.6	HD Aspect Ratio.....	4
2.6	Field/Frame Rate and Dominance	4
2.6.1	Field Rate.....	4
2.6.2	Field Dominance.....	4
3	Audio	5
3.1	Main Service Audio.....	5
3.2	Secondary Service audio.....	5
3.3	Audio Loudness Requirements.....	5
3.3.1	Short Form Content.....	5
3.3.2	Long Form Content.....	5
3.3.3	Pure Music Content.....	5
3.4	Average dialogue	6

3.5	Surround Sound Downmix Loudness.....	6
3.6	Audio Quality.....	6
3.7	Audio Phasing & Synchronization.....	6
3.7.1	Audio Phase.....	6
3.7.2	Audio Synchronization	6
3.7.3	Audio Synchronization between Services	7
3.8	Media Track Assignments for OP1A	7
3.9	Stereo Synthesis	7
3.10	Up-mixing	7
4	Ancillary Information	8
4.1	Ancillary Data	8
4.2	Time Code.....	8
4.3	Time Code Start Time	8
4.4	Closed Captioning.....	8
4.4.1	Closed Caption Data.....	8
4.4.2	Closed Captions Accuracy	9
4.4.3	Closed Captions Synchronization	9
4.4.4	Closed Captions Placement.....	9
4.4.5	Closed Captions Completeness	9
4.5	Active Format Description (AFD)	10
4.5.1	AFD Encoding	10
4.5.2	AFD Codes	10
5	Technical Specification for Containerization.....	11
5.1	Media Inventory	11
5.2	Containerized Submissions.....	11
6	File Submission	12
6.1	File Guidelines	12
6.2	File Nomenclature	12
6.3	File Contents.....	12
6.4	File Audio.....	12
6.4.1	Stereo Program	12
6.4.2	5.1 Surround Program.....	12
6.5	File Formats	13
6.5.1	XDCAM 422	13
6.5.2	DNx145, DNx220 or DNx220x	13
6.6	Feature: Captioning	14
6.7	Timecode track.....	14

7	Reference Table	15
A.1	Loudness Measurement of Long Form Content (“Programs”).....	17
A.1.1	Manually Isolating Dialogue	17
A.1.2	Speech-Gated Measurement	17
A.1.3	Live Production.....	17
A.2	Audio Levels between Programs	17
A.3	Music and Sound-Effects	18
A.4	Surround Sound	18
A.5	PBS Recommended Metadata Used in Emulation.....	20

1 Scope and Purpose

1.1 Purpose

This Technical Operating Specification, TOS, provides standards for producing programs of a consistently high technical quality for delivery to PBS.

1.2 Requirements

Organizations submitting programs to PBS must have an understanding of the technical specifications set forth in this document and proficiency at using professional digital measurement equipment.

1.3 Producer Responsibilities

Evaluate the program to meet these specifications prior to submission. Program submissions not meeting these specifications may be rejected requiring corrective action and delay in processing.¹

PBS accepts only HD programs submitted via file delivery (see section 6 File Submission).

1.4 File Based Media

- XDCAM 422 50Mb/s,
- DNxHD 145, 220 or 220x OP1a file

¹ Note: For questions related to program submissions and this TOS contact the Media Operations Center (MOC) at PBS MOOperations@pbs.org or (703) 739-5205.

2 Video

2.1 Video Image Quality

This section concerns the video image quality of submitted programs. All programs must be produced with modern digital acquisition and editing systems, with careful attention as early in the process as possible and beginning with the highest quality available, then maintaining high quality throughout the acquisition and post-production processes.

2.2 Video Definition

For the purpose of this document, HD and SD upconverted video will be defined as follows:

2.2.1 HD Image

The HD image format shall be 1920 X 1080 pixels, 4:2:2 with 8 or 10-bit sampling and compliant with SMPTE ST 274-2008 standard. These image specifications should be preserved as much as possible throughout the production process.

2.2.2 SD Upconvert

SD upconverted submissions are defined as those programs created in standard definition and then upconverted to high definition before submission.

2.3 Video Anomalies

Except in the case of use of archival content where no better copies are available, the image must be free of compression and analog artifacts.

2.3.1 Compression Artifacts

The image must be free of compression artifacts (such as macroblocking and mosquito noise), aliasing (such as the artifacts associated with scan conversion), frame dropouts and other artifacts associated with conversion and encoding.

2.3.2 Analog Artifacts

The image must be free of picture impairments associated with legacy analog equipment such as lag, smear, scratches, videotape dropouts, head switching, or composite video artifacts.

2.3.3 Non-Image Content

Only visual content intended to be seen by the viewer is allowed in the active image area. Particular care must be exercised to remove miscellaneous non-image content such as closed captioning information, particularly in images that have letter/pillar-boxed content.

2.4 Video Levels

2.4.1 Conversion

This specification requires Y' Cr' Cb' digital component submissions. Producers should be aware that 75% of Y' Cr' Cb' values do not exist in RGB color space and will be clipped by downstream consumer equipment before display.

2.4.2 Color Gamut

To ensure the consumer display is faithful to the submission, Y' Cr' Cb' values must fall within the equivalent RGB colorspace, with the exception of overshoots due to measurement resampling filters.

<i>Measurement Method</i>	<i>Valid Y & RGB Range</i>
8 bit Digital	16-235
IRE	0-100
Analog Equivalent	0-700 mV

Producers should be aware that the RGB gamut will be hard limited when the final PBS package is created for distribution.

2.4.3 Black Level

Objectionable white or black clipping must not be evident. Program black level must be set to 0v on the Y' waveform.

2.5 Video Image

2.5.1 Image Format

The image format shall be 1920 X 1080 pixels and compliant with SMPTE ST 274-2008.

2.5.2 Pillarbox Content

When 4:3 content is placed within a 16:9 production aperture (which is termed "pillarbox") the image will fill the following pixels on every line; thus horizontally centering the image (+/- 5pixels).

Table 1 4:3 Content in 16:9 Image

Format	Line Pixels
1080i 16X9 Pillarbox	241-1680

2.5.3 4:3 Protect

When content in a 16:9 production aperture is 4:3 protected (AFD code 1111) then the image will be centered in the middle 75% of the picture aperture; thus horizontally centering the image.

2.5.4 Safe Areas

Safe Action is the area within which all significant action must be contained as defined by SMPTE standard 2046-1. The area is defined as 93% of the width and height of the production aperture.

Safe Title Area is the area within which all significant text must be contained as defined by SMPTE standard 2046-1. The area is 90% of the width and height of the production aperture.

PBS Safe Area is the area located in the lower right. It is reserved for member stations to place their own station bug. A station bug will overlap with program IDs should the text extend into this area and therefore significant text or graphics entering this area is prohibited.

A graphic overlay that outlines this area can be downloaded at a link in the reference section.

2.5.5 Subtitles

Subtitles are limited to the Safe Title Area of 16X9 or 90% of the width and height of a 4:3 protected area (when a 4:3 protected area is present).

2.5.6 HD Aspect Ratio

The aspect ratio of all HD programming shall be 16:9. All video content submitted originally produced with an aspect ratio other than 16:9 (e.g. UK origin 14:9, theatrical widescreen releases, legacy 4:3, etc.) shall align within the 16:9 frame so that no black bars completely surrounding the active video (Postage Stamp effect) shall be observed.

2.6 Field/Frame Rate and Dominance

2.6.1 Field Rate

The field rate for high definition (1080i) format is 59.94 fields per second (60 multiplied by 1000/1001). The current frame rate for PBS high definition (1080i) format is 29.97 frames per second interlaced. Content submitted as 29.97 frame segmented (30 multiplied by 1000/1001) should be checked for temporal artifacts such as anomalies that might be present in horizontal crawls or vertical credit rolls when displayed at 1080i prior to submission.

2.6.2 Field Dominance

All 1080i files submitted to PBS must be created with field 1 (also referred to as top or odd) being dominant (first). There can be no changes in field dominance within any submitted media.

3 Audio

3.1 Main Service Audio

Main service audio is defined as the primary service that is intended to serve the majority of the audience. All Main Service Audio shall be stereo or 5.1 surround sound. Legacy material may contain mono audio if it was not originally captured in stereo. Only stereo will be accepted for Promos and interstitials.

3.2 Secondary Service audio

Secondary services include alternate languages and/or Descriptive Video Information (DVI).

3.3 Audio Loudness Requirements

Refer to ATSC A/85:2013, "Techniques for Establishing and Maintaining Audio Loudness for Digital Television" to understand PBS's requirements for loudness management and surround- sound management. This document is available at a link in the reference section. ¹

3.3.1 Short Form Content

As stipulated in Annex J of ATSC A/85, metering of short form content, (defined for PBS as 90 seconds or less), must conform to ITU-R BS.1770-3 for loudness measurement and true-peak measurement and must apply to all channels of 5.1-channel submissions (with the exception of the LFE channel which is not included in the loudness measurement), or to both channels of a 2-channel stereo program. For such short form content (such as underwriting spots, video offers, etc.), measured loudness shall be averaged over the duration of the body of the program, and the resultant average must conform to -24 LKFS, +/- 2 LU. True Peak levels shall not exceed -2 dBTP at any time.

3.3.2 Long Form Content

The audio mix of long form ("program") content must conform to the guidelines presented in ATSC A/85:2013, using ITU-R BS.1770-3 for loudness measurement and true-peak measurement and must apply to all channels of 5.1-channel submissions (with the exception of the LFE channel which is not included in the loudness measurement), or to both channels of a 2-channel stereo program. Measured loudness shall indicate the average loudness of normally spoken dialogue during the body of the program, and the resultant average must conform to -24 LKFS, +/- 2 LU. True Peak levels shall not exceed -2 dBTP at any time. See Appendix A for guidance in measuring loudness.

- The loudness of normally spoken dialogue must remain consistent from act to act and throughout the body of the program.
- The loudness of music and sound effects must not overwhelm, but rather it must be creatively blended with the loudness of dialogue in such a way that most viewers will easily discern the dialogue on first listening.

3.3.3 Pure Music Content

In some programs, such as music performance programs where the length of dialogue is very short and the remainder of the program audio is not dialogue, the anchor element becomes the music.

¹ Producers should be aware that secondary audio loudness for mono/DVI/SAP audio channels is 3 dB lower than the main service audio. Secondary audio should be delivered to PBS in balance with the main service audio and any adjustments to correct this loudness must be performed downstream from PBS. To correctly adjust to a -24 LKFS +/-2 audio loudness the mono/DVI/SAP channels should either be increased by 3 dB during playback or the emission encoder dialnorm should be changed from -24 to -27.

- The loudness measurement of such non-dialogue main portions of these programs is to be integrated (averaged) over the “body of the program”, using BS.1770-3 LKFS meters.
- The target loudness of the “body of the program” is to be -24 LKFS +/- 2 LU when AVERAGED for the duration of the program.¹
- The loudness of any dialogue (e.g., introductory comments, promos, etc.) in such music performance programs also must match the average loudness level of -24 LKFS +/- 2 LU.

3.4 Average dialogue

Programs may have music or effects true-peak levels no higher than -2 dBTP during moments of dramatic impact, as long as average dialogue levels are maintained as specified here.²

3.5 Surround Sound Downmix Loudness

Surround-sound programs shall be measured for downmix loudness compatibility using downmix settings of -3 dB in the center channel and -3 dB in the surround left and right channels. The audio mix of 5.1-channel programs must be created in such a manner that the stereo downmix shall have overall loudness levels of -24 LKFS +/- 2 LU. This generally entails not spreading the center image across the front channels. The Low Frequency Effects (LFE) channel is not included in the BS.1770 algorithm.

3.6 Audio Quality

- 3.6.1 The audio mix must be free of audible clipping and other distortions.
- 3.6.2 The audio mix must be free of objectionable noise such as loud hiss.

3.7 Audio Phasing & Synchronization

3.7.1 Audio Phase

To insure compatible stereo and mono down-mix, all channels must be in phase within an audio service.

3.7.2 Audio Synchronization

Main service audio must not lead or lag video by more than one-half frame as illustrated in Figure A: Audio Sync. This specification applies to all program elements.

Figure A: Audio Sync



¹ It is understood that music with wide dynamic range could fall outside this specification for periods of time within the program body.

² Producers lacking true-peak measurement tools should be aware that IEC peak measurements often result in readings approximately 2 dB lower than true peak. Producers should also be aware that some legacy downstream stations or MSO systems may clip signals at levels higher than -10 dBFS.

3.7.3 Audio Synchronization between Services

Audio sync must be within one frame between services (e.g., between main and descriptive video information, or music and effects between main and alternate language).

3.7.4 Media Track Assignments for OP1A

3.8 Stereo

High definition programs delivered in stereo via an OP1a file format must have discrete audio channels with the following track assignments.

Table 4 Media Track Assignments for File Based Content

Channel/Track	Assignment
Channel 1	Left Stereo (L)
Channel 2	Right Stereo (R)
Channel 3	Silence
Channel 4	Silence
Channel 5	Silence
Channel 6	Silence
Channel 7	DVI (Mono) or Mono Mix
Channel 8	Alternate Language (Mono) or Mono Mix

3.9 5.1

High definition programs delivered in 5.1 surround sound via an OP1a file format must have discrete 5.1 audio channels with the following track assignments.

Table 4 Media Track Assignments for File Based Content

Channel/Track	Assignment
Channel 1	Left Front (L)
Channel 2	Right Front (R)
Channel 3	Center (C)
Channel 4	Low Frequency Effects (LFE)
Channel 5	Left Surround (Ls)
Channel 6	Right Surround (Rs)
Channel 7	DVI (Mono) or Mono Mix
Channel 8	Alternate Language (Mono) or Mono Mix

3.10 Stereo Synthesis

Stereo synthesizing is not allowed within any content at any time.

3.11 Up-mixing

When up-mixing two channels (stereo) to multi-channel surround sound (5.1 channel), audio must be properly distributed among the channels. The resulting center image should not be spread from center. Up-mixed audio must be downmix-compatible to stereo and mono, that is, the resultant downmix should be virtually indistinguishable from the original.

4 Ancillary Information

4.1 Ancillary Data

The ancillary data must be limited to:

- 436M EIA 608/708 captions DID 61, SDID 01 for CEA708 and DID 61, SDID 02 for compatibility bytes may be included
- SMPTE 2016-3 compliant Active Format Description (AFD) data may be included DID 41 SDID 05

4.2 Time Code

Drop frame timecode on DID 64, SDID 07 and LTC DID 64, SDID 64

4.3 Time Code Start Time

The recording must begin with time code starting at 00:59:45;00, with 01:00:00;00 being the start of program material. The time code must increment without interruption from the beginning of the file until 15 seconds after the final program segment.

4.4 Closed Captioning¹

PBS strives to ensure that all viewers have full access to its video services and programs. PBS specifications require that viewers who are deaf or hard of hearing be provided closed captions that are accurate, synchronized, properly placed and complete, Pursuant to the FCC's Report and Order, Declaratory Ruling, and Further Notice of Proposed Rulemaking, (FCC 14-12 Feb. 24, 2014), PBS adopts the following non-technical quality standards for all closed captioning:

4.4.1 Closed Caption Data

- Closed Caption data must be encoded as specified in CTA 708 with 608 compatibility bytes
- The primary language captioning data stream must be carried in the Primary Synchronous Caption Service (708 Service 1 / 608 CC1).
- The optional secondary captioning service must be carried in the Secondary Synchronous Caption Service (708 Service 2 / 608 CC3)
- ~~Both 608 & 708 closed captioning data are required at time of delivery. MPEG2 user data is not allowed.~~

¹ Closed Captioning is not required for interstitial material, promotional announcements, and public service announcements that are 10 minutes or less in duration and outside the body of the program.

4.4.2 Closed Captions Accuracy

- Captions must match the words spoken in the dialogue.
- Captions must be in the original language (i.e., English or Spanish), to the fullest extent possible.
- Captions must include full lyrics of songs that are included in the audio track.
- Captions must contain all words, in their spoken order, without paraphrasing or other substitution.
- Spelling must be correct.
- There must be appropriate punctuation and capitalization to reflect the flow of the dialogue.
- Correct verb tenses must be used.
- There must be accurate representation of numbers, including related symbols such as dollar signs, etc.
- It is not permitted to rewrite dialogue, including slang or grammatical errors in dialogue.
- Utterances such as “um” and other irregularities such as false starts, etc., need to be mirrored by the captions.
- Synonyms may not be used to replace words of dialogue.
- Best efforts must be made to convey the manner and tone of the speaker’s voice.
- Other non-verbal audio cues which are not visibly obvious must be conveyed by captions to the fullest extent possible.
- Captions must include an indication of who is speaking, the existence of music, sound effects, and audience reaction to the greatest extent possible, given the nature of the program.
- If there is more than one speaker, the captions must identify each speaker through caption identification or caption placement, making it clear who is speaking at any given time.
- When speakers are off-screen, captions must identify who is speaking if hearing viewers are able to discern who is speaking.
- Captions must be legible, with proper spacing between words.
- Captioning must be free from missing, garbled or extraneous letters.

4.4.3 Closed Captions Synchronization

- Captions must coincide with their corresponding spoken words and other sounds to the greatest extent possible, given the nature of the program.
- Captions are to begin to appear at the time the corresponding speech or sounds begin.
- Captions should end when the speech or sound ends.
- Captions are to display on the screen at a speed that can be read by viewers.

4.4.4 Closed Captions Placement

- Captions shall not block important visual content on the screen, including (but not limited to) character faces, featured text such as weather, news updates, graphics, credits, etc., and other information, that is important for understanding program content.
- Proper placement requires that caption fonts be sized properly for legibility.
- Lines of caption shall not overlap one another.
- Captions must not run off the edge of the screen.

4.4.5 Closed Captions Completeness

- Captions must run from the beginning of the program to the end of the program to the fullest extent possible.
- Captions must be complete for the entirety of the program including clear caption marks at the end in order for a program to be fully accessible.

4.5 Active Format Description (AFD)

Inclusion of AFD data on media submissions is voluntary at the present time. If it is submitted it shall conform to the following specifications.

4.5.1 AFD Encoding

An AFD code must be placed on line 11 of every field of every frame of each segment including the pre- and post-segment black portions. If AFD codes are present in both ancillary data and encoded video data then the ancillary data takes precedence.

4.5.2 AFD Codes

Only the following 16:9 AFD codes may be used¹

- AFD code 1000 (8) -- Full frame 16:9 image
- AFD code 1001 (9) -- 4:3 pillarbox image centered in frame
- AFD code 1010 (10) -- Full frame 16:9 image – All image area is protected, cropping is not allowed
- AFD code 1111 (15) -- 16:9 image with alternate 4:3 center – Visual information outside 4:3 protected area may be cropped with minimum impact for the viewer.

¹ Note: Currently AFD data will be overwritten during the PBS containerization process using the AFD Code submitted in the Orion Media inventory submission for each segment.

5 Technical Specification for Containerization

5.1 Media Inventory

A container is a cut as identified by the media inventory form which could be either short form or long form.

5.2 Containerized Submissions

- The first program segment must always start at 1:00:00;00 timecode.
- Unless a black segment is defined as “not for air” it shall be considered part of the container. Captioning data must be completely contained within each container including clear caption marks at the end.
- Alternate Language and Descriptive Video Information must be contained completely within the container so that the container can stand alone regardless of the adjacent material.
- Audio loudness on each container should meet audio loudness specifications for long or short form content. Please reference Section 4 for full audio loudness specifications.

6 File Submission

6.1 File Guidelines

- Producer must follow the PBS guidelines for file delivery to qualify. Contact PBS for instructions on the file delivery guidelines.
- A backup file must also be sent via an approved alternate delivery method for any program submitted via file delivery less than 3 business days to air.

6.2 File Nomenclature

Files will be named using the media ID provided by PBS Packaging Forms.

6.3 File Contents

- Files will contain a minimum of 15 seconds at the beginning; 15 seconds of clean black or slate information identifying the included media followed by 2 seconds of clean black and 15 seconds of clean black run-out at the end of the file. Files will contain a time code track that provides continuous, sequential timecode from start to end of file. Timecode should start at 00:59:45;00 with the first active program material starting at 01:00:00;00.

6.4 File Audio

Files will have audio configured in one of the following:

6.4.1 Stereo Program

- Chan 1 Left Stereo
- Chan 2 Right Stereo
- Chan 3 Silence
- Chan 4 Silence
- Chan 5 Silence
- Chan 6 Silence
- Chan 7 Mono mix or DVI
- Chan 8 Mono mix or SAP (secondary language)

6.4.2 5.1 Surround Program

- Chan 1 Left Front
- Chan 2 Right Front
- Chan 3 Center
- Chan 4 LFE
- Chan 5 Left Surround
- Chan 6 Right Surround
- Chan 7 Mono mix or DVI
- Chan 8 Mono mix or SAP (secondary language)

6.5 File Formats

6.5.1 XDCAM 422

XDCAM 422 File must meet Sony XDCAM file specifications including File Container: MXF OP1a

6.5.1.1 Video:

- 1920x1080i
- 29.97 fps
- 4:2:2
- MPEG-2 Long GOP @ 50 Mbps

6.5.1.2 Audio:

- 8 channels
- 24 Bit
- 48 KHz sampling rate
- 1152 kbps Bitrate

6.5.1.3 Ancillary data:²

- 436M 708 captions with 608 compatibility bytes
- Timecode track:
- Drop frame timecode

6.5.1.4 Timecode track:

Drop frame timecode on DID 64, SDID 07 and LTC DID 64, SDID 64

6.5.2 DNx145, DNx220 or DNx220x

DNx145, DNx220 or DNx220x files must meet file specifications including File Container: MXF OP1a

6.5.2.1 Video:

- 1920x1080i
- 29.97 fps
- 4:2:2
- DNxHD @ 145 or 220 Mbps

6.5.2.2 Audio:

- 8 channels
- 24 Bit
- 48 KHz sampling rate
- 1152 kbps Bitrate (shown in file metadata)

6.5.2.3 Ancillary data:

- 436M 708 captions with 608 compatibility bytes
- Timecode track:
- Drop frame timecode

6.5.2.4 Timecode track:

Drop frame timecode on DID 64, SDID 07 and LTC DID 64, SDID 64

² 436M EIA 608/708 captions DID 61, SDID 01 for CEA708 and DID 61, SDID 02 for compatibility bytes may be included. SMPTE 2016-3 compliant Active Format Description (AFD) data may be included DID 41 SDID 05
PBS Technical Operating Specification 2020 Part 1

6.6 Feature: Captioning

- All media must have accompanying captions on SDID 01 for CEA708 and DID 61, SDID 02 for compatibility bytes or a .scc or SMPTE Timed Text xml caption file.³
- The caption file should include an EDM (Erase Displayed Memory command), also known as a clear pulse, at the first frame matching video and also at the beginning and end of each container. Captioning must accompany file submissions.
- SCC formatted Captioning file (.scc extension)
- SMPTE Timed Text Captions ST 2052-1:2010 - Timed Text Format ⁴

6.7 Timecode track

Matching timecode of media audio is to be associated with Drop Frame Time Code

³ If a program is submitted with embedded captioning and a sidecar caption file, the sidecar file will overwrite the embedded captioning upon processing at PBS. Note: Tunneled 608 data should not be included in Timed Text file.

7 Reference Table

References are reviewed on a regular basis to keep up to date with the respective versions. SMPTE references are subscription based and only the abstract description is available to view without paid subscription.

#	Item	Description
1.	ST-274	ST 274:2008 - For Television 1920 × 1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequences for Multiple Picture Rates http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7290129&contentType=Standards
2.	ST-2046	ST 2046-1:2009 - Specifications for Safe Action and Safe Title Areas for Television http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7291650&contentType=Standards
3.	A-85	ATSC Recommended Practice: Techniques for Establishing and Maintaining Audio Loudness for Digital Television (A/85:2013) http://atsc.org/wp-content/uploads/2015/03/Techniques-for-establishing-and-maintaining-audio-loudness.pdf
4.	ST-334	ST 334-2:2007 - Caption Distribution Packet (CDP) Definition http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7290879&contentType=Standards
5.	ST-2016	ST 2016-3:2009 - Vertical Ancillary Data Mapping of Active Format Description and Bar Data http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7290864&contentType=Standards
6.	EG-40	EG 40:2002 - Conversion of Time Values between SMPTE 12M Time Code, MPEG-2 PCR Time Base and Absolute Time http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7291287&contentType=Standards
7.	FCC-R&O	Closed Captioning Quality Report and Order, Declaratory Ruling http://www.fcc.gov/document/closed-captioning-quality-report-and-order-declaratory-ruling-fnprm
8.	CTA-708	Digital Television (DTV) Closed Captioning: Document Number: CTA-708-E (ANSI) https://members.cta.tech/ctaPublicationDetails/?id=a10017f5-a01c-e811-90ce-000d3a0072d6

		Subcommittee/CEA-708-D.aspx
9.	RP-219	RP 219:2002 - High-Definition, Standard-Definition Compatible Color Bar Signal http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7289865&contentType=Standards
10.	ST-2052	ST 2052-1:2010 - Timed Text Format (SMPTE-TT) http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7292145&contentType=Standards
11.	ATSC A/85:2013	https://www.atsc.org/wp-content/uploads/2015/03/Techniques-for-establishing-and-maintaining-audio-loudness.pdf
12.	FCC-14-12A1	https://docs.fcc.gov/public/attachments/FCC-14-12A1.pdf
13.	Rec. 709	https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.709-6-201506-I!!PDF-E.pdf

Appendix A. Audio Recommended Practices

In all cases, to enable correct judgments of loudness, imaging, and other issues of quality-control, all content should be monitored on properly aligned and calibrated loudspeakers as per ATSC A/85. This means aligning the 5.1-channel loudspeaker geometry to the ITU-R BS.775 standard and calibrating their loudness so that normally spoken dialogue at -24 LKFS loudness will reproduce at 78 dB SPL in most broadcast-type spaces (depending on the room volume, as explained in A/85). Monitoring environments in remote trucks should be calibrated to 76 dB SPL, as per A/85.

A.1 Loudness Measurement of Long Form Content (“Programs”)

As stated in 3.1, loudness measurements of normal long form program content are to be taken with systems employing the BS.1770-3 algorithm. There are two broad categories of loudness measurement techniques that can yield acceptable results, as follows:

A.1.1 Manually Isolating Dialogue

When given loudspeaker monitors that have been properly aligned and calibrated as per A/85, skilled mixers and other skilled operators can easily recognize normally spoken dialogue by listening. In post-production, it can be an easy matter to manually isolate portions of content representing such dialogue which, assuming consistency of levels, may serve as proxies for the overall program loudness. However, it will remain true that more consistently reliable results will be obtained when more sections of representative dialogue are included in the cumulative (“integrated”) measurement of the average loudness.

A.1.2 Speech-Gated Measurement

Very consistent results can be obtained by using an automated speech-gating algorithm for controlling the loudness measurement. In common usage, this may be known as “Dialogue Intelligence”. Such systems automatically identify normally spoken dialogue and use that for controlling (“gating”) the cumulative averaging (“integrating”) of the loudness measurement. In situations where selecting Dialogue Intelligence on device may force the selection of the BS.1770-1 algorithm, doing so is acceptable.

A.1.3 Live Production

In live production, the first and most important line of defense is in having properly aligned and calibrated loudspeaker monitors, as per A/85. When given that, most mixers can immediately tell by listening when normally spoken dialogue is at the proper loudness. The reliability of such judgments is enhanced by rehearsing with normalized dialogue. Generally, confidence will be enhanced by at least occasionally spot-checking such results during the production with quick glances at a measurement device employing the BS.1770-3 algorithm, especially when the measurements are presented to the mixer in a way that is easily and quickly read, such as a color-coded bar-graph representation.

A.2 Audio Levels between Programs

Audio levels within and between containers and their transitions must be consistent. Consumer dissatisfaction arises when there are abrupt level changes. Therefore, even when the overall average loudness of a program conforms to the standard, special attention should be paid to short term loudness variations at the end of programs so as to ensure a smooth transition to the content that will follow. It is strongly recommended to ensure that such short-term dialogue loudness variations are within the Comfort Zone presented in A/85, as shown in Figure B, below.

A.3 Music and Sound-Effects

Music and sound-effects loudness levels must be creatively blended with dialogue loudness so that the dialogue is dominant as the anchor element to ensure that a wide variety of viewers can understand the dialogue upon first viewing, including when in home listening conditions with high ambient noise, moderate program levels, and a wide variety of possible audio monitoring systems.

A.4 Surround Sound

In the interest of presenting a broad range of desirable listening positions within a consumer 5.1-channel listening environment (where not only should the illusion of the dialogue appear to emanate from the picture, but likewise the same is also true of the center imaging of music and sound-effects, including ambience, even when listening off the centerline), it is recommended that all center imaging (not just dialogue) be sent to the center channel of a 5.1-channel mix. When feasible, it is suggested that this be actively accomplished by the mixer as part of the creative mixing process. When that is not possible, acceptable results can often be subsequently accomplished with an automatic up-mixing system. Up-mixing stereo content can create a 5.1-channel mix with a strong center channel, as well as immersive content in the rear channels.⁵

It is highly recommended that all center image content (including most dialogue) should not rely on phantom imaging to create the illusion of a center image, but rather it should be hard assigned to the center channel of 5.1-channel surround-sound programs, except as necessary for dramatic content to be consistent with camera perspective. Thus, when listeners are positioned off the centerline, the illusion of the center audio image should not appear to shift away from the center of the picture.

Center image content should not be spread across the front channels. Doing so would adversely affect the subsequent downmix loudness, and would also adversely affect the quality of experience presented to home viewers.

In shows produced with live audiences, it is recommended that the auditory image of the audience should relate to the camera perspective. This usually means that the auditory image of the audience should not be mixed to the front. Usually camera perspective will suggest that the image of the audience should be more immersive by being mixed predominantly to the sides and

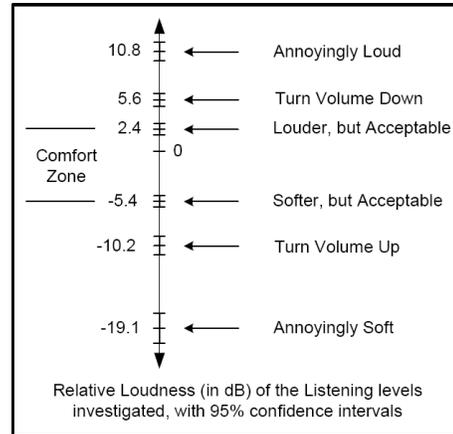


Figure B

¹Note: Producers should be aware that programs submitted in 5.1-channel audio will be heard by the majority of consumers as a stereo downmix. This downmix will occur at stations without 5.1 transmission equipment, within cable MSO receivers for analog retransmission, and in consumer stereo AC3 decoders.

rear. Usually, the goal should be to create the illusion that the viewer is actually sitting in the audience, rather than having the auditory suggestion that the audience is in front of the viewer, unless for some reason the camera perspective suggests such a thing.

It is important for content producers to listen to an accurate representation of what will be experienced by the viewing/listening audience. In this regard, it is recommended to use a system of emulation¹ which precisely emulates the listening experience of the home viewers (when programmed with the same metadata as will ultimately be decoded by home receivers). This will enable not only things such as listening to and measuring the loudness of the metadata-derived stereo (and mono) downmix of surround-sound content, but also other things such as listening to and visualizing the dynamic range control that is usually (but optionally) applied by the DRC (Dynamic Range Control) feature of ATSC home receivers.

Proper emulation relies on using the same metadata that will be presented to the audience. PBS recommends using the metadata values shown in the table at the end of this section. This is the metadata that is expected to be encoded by the member stations at the time of their AC-3 encode, and thus is the same metadata that should be used in emulation.

¹ The Dolby DP570 and Media Emulator are examples of such devices that can precisely emulate the home listening experience. If an emulator is unavailable, a similar experience can be had by passing the content through the AC-3 encode cycle (where metadata will be created) and then through the decode cycle, e.g. by a consumer receiver, where the metadata will be applied to the content that is heard, but without some of the other valuable information provided by an emulator, such as the visible display of the compression and expansion action of the DRC system.

A.5 PBS Recommended Metadata Used in Emulation during Production/Post-Production

Table 6 Audio Meta Data Parameters

METADATA PARAMETER	METADATA VALUE
Dialogue Level (dialnorm)	-24 dB
Channel Mode (5.1-channel surround)	3/2
LFE Channel	Enabled
Data Rate	448 kbps
Bitstream Mode	Main Complete
Line Mode Profile	Film Standard
RF Mode Profile	Film Standard
RF Overmodulation Protection	Off
Center Downmix Level	-3 dB
Surround Downmix Level	-3 dB
Dolby Surround	Not Indicated
Copyright Bit	Yes
Original bitstream	Yes
Audio Production Information	None
Mixing Level	78 dB SPL
Room Type	Not Indicated
Extended Bitstream	Disabled
Preferred Downmix	Lo/Ro
Dolby Surround EX	Not Dolby EX
De-emphasis	Autodetect
DC Filter	Enabled
Lowpass Filter	Enabled
LFE Lowpass Filter	Enabled
Surround 3 dB Attenuation	Disabled
Surround Phase Shift	Disabled
Input Channel Assignment	L/R C/LFE Ls/Rs
Clock Source	Digital Input
Input Format	PCM
Metadata Source	Internal
Reversion Mode	Last Used