TANDBERG television

White Paper

Preparing for the Broadcast Analog Television Turn-Off: How to Keep Cable Subscribers' TVs from Going Dark

Learn about the NTSC-to-ATSC converter/ receiver and how TANDBERG Television's RX8320 ATSC Broadcast Receiver solves the Analog Turn-Off issues discussed in this paper

Matthew Goldman
Vice President of Technology
TANDBERG Television, part of the Ericsson Group

Table of Contents

1. The "Great Analog Television Turn-Off"

This is history in the making: After 68 years of over-the-air television broadcasting, analog television signals (in accordance with the National Television System Committee (NTSC) standard) will cease to exist in the United States on February 17, 2009. On this date, all full power stations in the U.S. will transmit only digital television (DTV) signals, either in high definition (HD) or standard definition (SD) format (per the Advanced Television Systems Committee (ATSC) standard). However, analog TV will continue for millions of cable subscribers. To ensure that Americans are not disenfranchised and unable to watch their local broadcast television channels, the Federal Communications Commission (FCC) is requiring cable TV operators to continue transmitting analog signals for a minimum of three years for all stations requesting must-carry status. Besides, with the potential confusion over the disappearance of one of America's favorite pastimes for certain demographics that have been relying on "rabbit ears" to receive analog TV, offering analog cable TV service post-analog broadcast turn-off may ironically become a boom to the cable TV industry.

Creating a replacement analog TV signal directly from a DTV signal is relatively easy; however, ensuring that all of the necessary ancillary and metadata – such as closed captions, picture aspect ratio, content advisory (formerly called program rating, and also referred to as V-chip technology), automated measurement of line-ups, and program guide data – is passed through the conversion process is not so straight forward. In fact, if the source is an HD picture with multi-channel surround sound, these formats also need to be converted to SD and stereo/mono sound in order for the more than 250 million analog TV sets in the U.S. to function properly (an estimated 15% of which are in households that receive analog TV over-the-air and not via a pay TV service, plus an estimated 30 million analog TV sets in pay TV households that are not connected to the pay TV service; approximately 65-70 million analog TV sets total [1]).

1.1 Receiving Over-the-Air TV Transmissions

Figure 1 shows a simplified example of the over-the-air front-end of a cable TV head-end today.

The local broadcasters' NTSC signals are received by NTSC receivers that demodulate the incoming analog broadcast signals to baseband and pass them to the analog plant for further processing. Similarly, ATSC receivers are used to demodulate the incoming DTV broadcast signals and pass the transport streams to the DTV plant for further processing.

Figure 2 shows a similar example of the over-the-air front-end of a cable TV head-end after February 17, 2009.
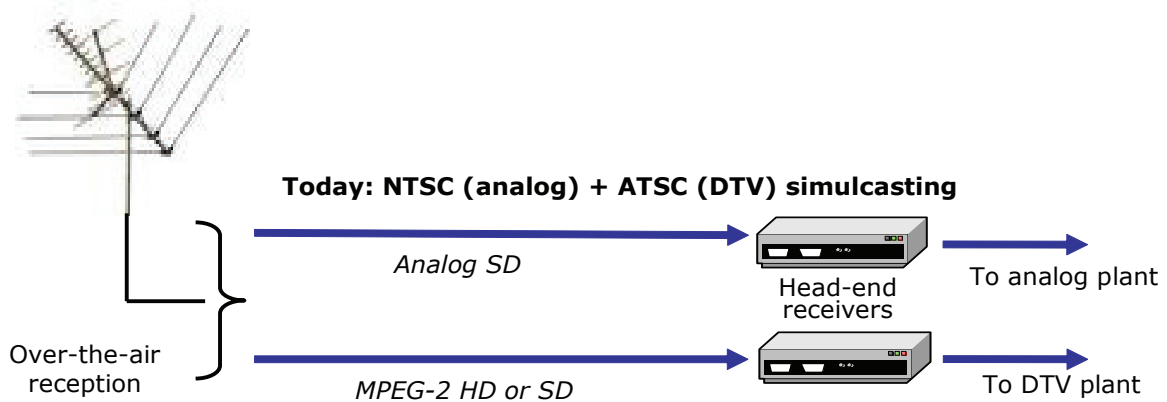

Figure 1. A simplified example of the front-end of a cable TV head-end today.
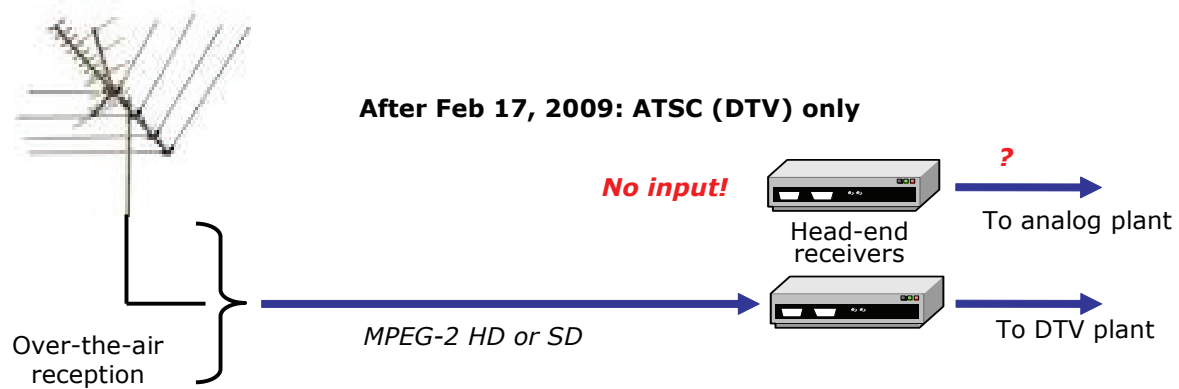
After Feb 17, 2009: ATSC (DTV) only

No input!

?

Head-end
receivers

To analog plant

To DTV plant

Over-the-air
reception

MPEG-2 HD or SD

Figure 2. A simplified example of the front-end of a cable TV head-end post broadcast analog turn-off.

As is clear in the figure, the analog plant no longer has analog TV signal sources because the local broadcast stations are no longer transmitting NTSC signals.

Some cable head-ends receive analog TV signals over a fiber interconnect (instead of or in additional to over-the-air reception). The head-end front-end is similar to Figure 2 above, except that the antenna is replaced with – or augmented with – a fiber interconnect. While analog TV signals can continue to be sent over fiber interconnects beyond the over-the-air turn-off date, in most cases the local broadcasters have indicated that they will not be continuing this practice and, in fact, a number of them are planning to dismantle their internal SD/analog plants.

Many local broadcasters also have announced that post analog turn-off they will not be simulcasting SD DTV versions of HD DTV sources. Consequently, comprehensive HD ATSC (DTV) to-NTSC (analog) conversion devices are needed in order to provide the necessary sources to feed the existing cable analog plant, such that no other changes need be made to the existing plant. Figure 3 shows an example of such a device.

After Feb 17, 2009: ATSC (DTV) only

New ATSC-to-NTSC
converter/receiver

Analog SD

To analog plant

MPEG-2 HD or SD

To DTV plant

Over-the-air
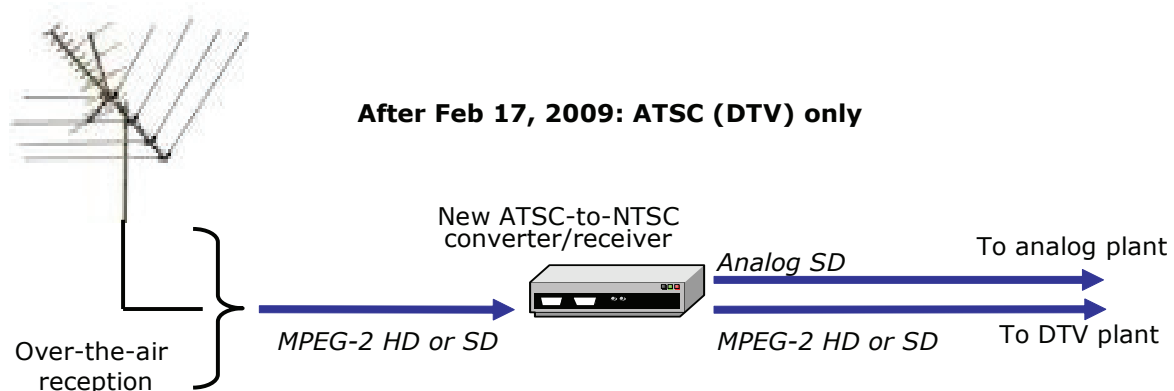reception

MPEG-2 HD or SD

Figure 3. The addition of an HD DTV-to-analog converter to the front-end of a cable TV head end post broadcast analog turn-off.

1.2 ATSC DTV to NTSC Analog Conversion

At a minimum, the new ATSC-to-NTSC converter/receiver needs to perform the following functions:

- Decode the MPEG-2 compressed video to baseband
- If the source image format is HD, down-convert to SD
  - Subsample the image
  - Perform picture aspect ratio conversion, 16:9 raster to 4:3 raster
  - Perform color space correction
- Decode the Dolby® Digital compressed audio to baseband
- Transform the primary multi-channel surround sound to stereo/mono
  - Apply program loudness adjustment by using dialog normalization
  - Perform dynamic range compression
  - If the number of source audio channels is greater than two, down-mix to stereo/mono
- Support an alternate audio service, either down-mix or pass secondary audio program (SAP) and/or video description service
- Extract DTV closed captions, if present, and convert to analog captions
- Extract embedded extended data services (XDS) information from the DTVCC Transport Channel, if present, and convert to analog XDS, including content advisory (V-chip)
- Extract other ancillary data carried in the incoming transport stream, if present, and convert to analog waveforms (e.g., Nielsen audience measurements and Macrovision Solutions Corporation TV Guide program data)

2. Video Down-Conversion

If the source video format is HD, the images must be down-converted to SD. This involves down-scaling or sub-sampling the picture resolution – basically throwing away image detail – as well as converting the picture aspect ratio from the 16:9 native raster of HDTV sets to the historical 4:3 native raster of NTSC analog TV se. Figure 4 shows the HD and SD scan formats overlaid on each other to emphasize the amount of down-scaling required to reduce the picture resolution to SD for analog TVs. Notice the more "square" 4:3 picture aspect ratio as opposed to the more rectangular 16:9 picture aspect ratio of HDTVs.
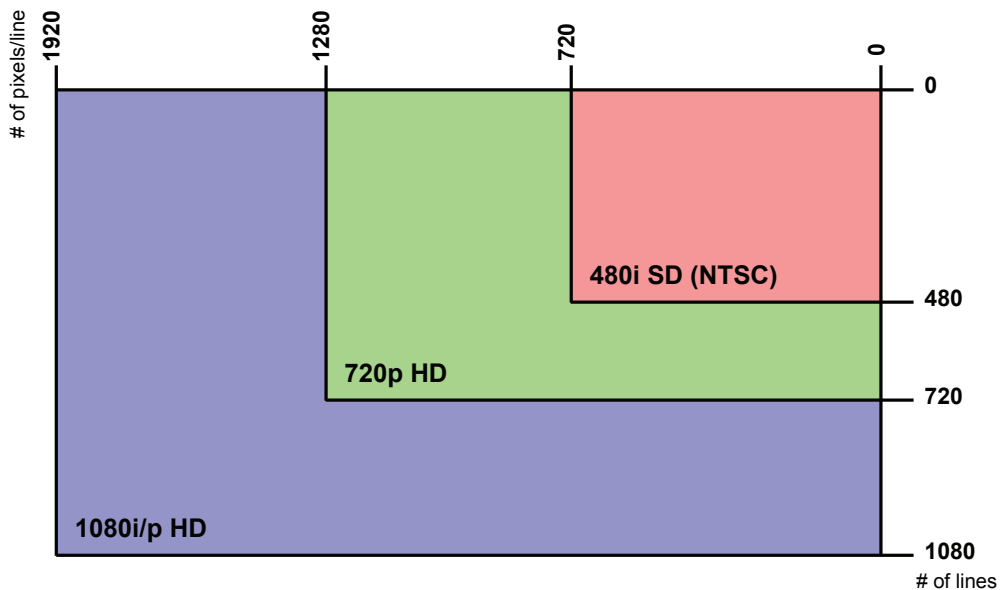

Figure 4. HD and SD picture resolutions.

In the down-conversion process, not all areas of the source HD picture may be of equal value. If the HD source was originally SD and through an up-conversion process the image was up-scaled, the left and right sides of the 16:9 image may not contain useful information. These images are often presented with black pillar bars, known as a ***pillarbox***, on a 16:9 HD raster (see Figure 5). If the HD source was produced in HD, but with SD down-conversion being considered

during production, then it may have been produced such that the *area of interest* is the 4:3 center of the 16:9 HD scan format, even though there is active information throughout the 16:9 image (see Figure 6). In both of these cases, the area of the 16:9 raster outside of the 4:3 center may be simply cropped to create the appropriate full-scale 4:3 aspect ratio of the analog TV. Although the HD source contains active information in the entire 16:9 raster, no critical information has been lost when the area outside of the 4:3 center is cropped. This process is commonly used and is known as *center-cut extraction*.

If the area of interest in the HD source is the entire 16:9 raster, then a center-cut extraction would compromise or "destroy" the viewing experience. In this case, the entire 16:9 image needs to be contained within the 4:3 SD raster. When this occurs, the top and bottom portion of the 4:3 display will not contain useful information, and black bars will appear at the top and bottom. This is referred to as a *letterbox* and is shown in Figure 7.
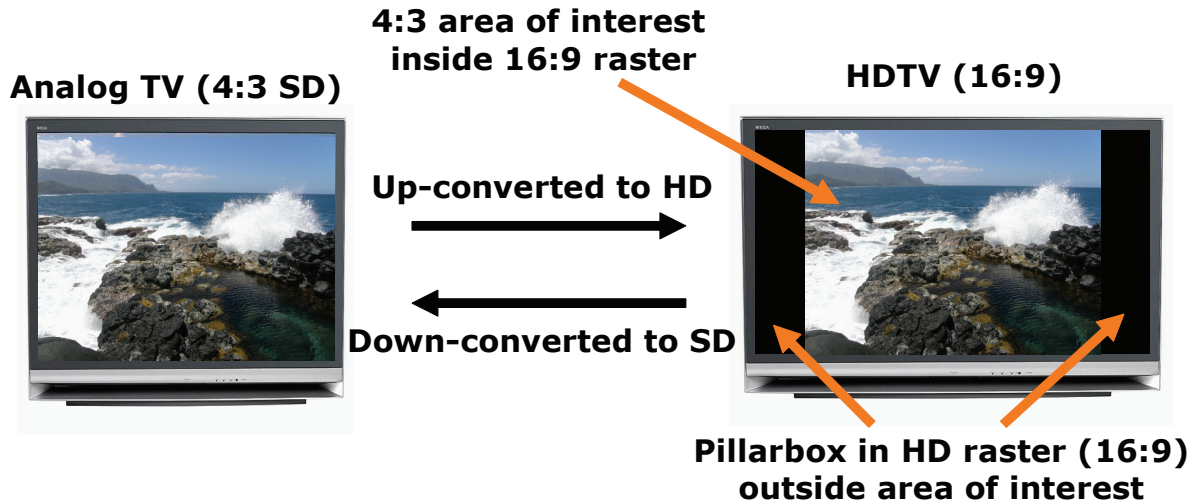

Figure 5. SD image up-converted to HD, resulting in a pillarbox to fill HD raster.
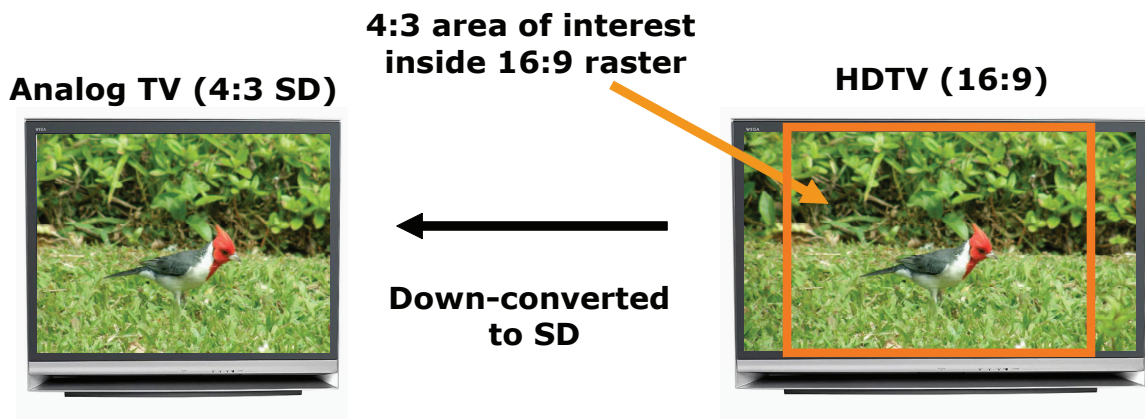*The reverse process is HD image center-cut extracted and down-converted to SD.*

Figure 6. HD image (full-raster) with shoot and protect 4:3 center, down-converted to SD.

Area of interest extends outside 4:3 center

Analog TV (4:3 SD)

HDTV (16:9)

Down-converted to SD
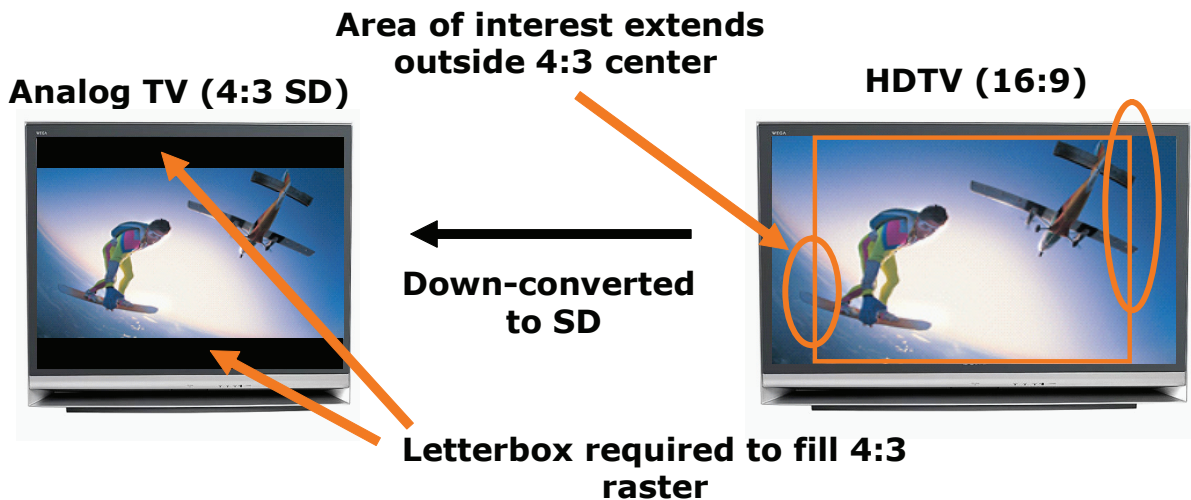
Letterbox required to fill 4:3 raster

Figure 7. HD image (full raster) down-converted to SD, resulting in letterbox to fill SD raster.

Unfortunately, there is no reliable way to determine the area of interest just from processing the image itself. As such, the HD-to-SD down-converter simply would be configured to either center-cut or letterbox (though more commonly center-cut). What happens if the wrong choice is made? The so-called *postage stamp* occurs, which is a significant source of complaints from the viewing audience. One of the common causes of postage stamping on an analog TV is when content that was originally SD has been up-converted to HD, and a pillarbox is added to fill the 16:9 HD raster (as previously shown in the HDTV on the right side of Figure 5). If this content is broadcast in HD, then there is no problem. If, however, the HD source is down-converted to SD somewhere further down in the processing flow – such as when the cable TV provider requires an SD version of the HD programming to support the analog tier – the down-converter only knows that an HD signal is present. To avoid possibly compromising the integrity of the content, the safe choice is to use the entire HD raster during down-conversion, and perform letterboxing to fill the SD raster (as previous shown in the analog TV on the left side of Figure 7). As shown in Figure 8, the end result is the postage stamp, a reduced size image in the raster, and a major annoyance to the viewing audience.
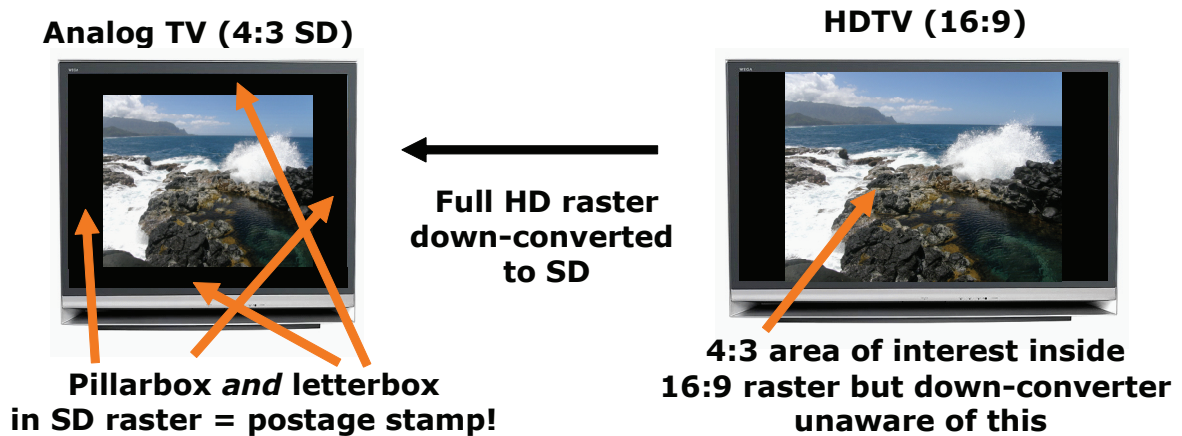

Analog TV (4:3 SD)

HDTV (16:9)

Full HD raster down-converted to SD

Pillarbox *and* letterbox in SD raster = postage stamp!

4:3 area of interest inside 16:9 raster but down-converter unaware of this

Figure 8. Down-conversion can cause the "dreaded" postage stamp, a reduced size image.

If the source content is carefully controlled, then a default configuration to down-convert could work effectively. However, more and more HD-only content is being produced today either without regard to ensuring that the area of interest is wholly contained within the center 4:3 area of the 16:9 raster, or for artistic reasons, the area of interest requires the full 16:9 raster. Local broadcasting stations also may have a desire to continue to offer viewers motion picture productions, which are typically produced in scan formats wider than 16:9 and, in many cases, can not be center-cut without compromising the integrity of the production. For these reasons, an automated, reliable method to determine the proper picture aspect ratio is much more desirable.

2.1 Active Format Description

Active format description (AFD) is a signaling method that may be used to determine automatically the proper picture aspect ratio to use during up-conversion and down-conversion processing. While the concept has been existence for several years, only recently was a standardized mechanism – SMPTE 2016-3 – created to pass this information to the DTV compression system located at the broadcast station [2]. The AFD indicates the area of interest. It is defined during production, and the signal processing equipment needs to interpret it in context to the function being performed. If the signal processing equipment alters the picture aspect ratio, then it may be necessary to also change the AFD passed on to the downstream equipment that will be processing the AFD in order to ensure the correct picture aspect ratio is obtained. In the case of down-conversion to analog TV, this is the terminus for AFD use.

AFD codes exist for both 4:3 and 16:9 scan formats [3]. Figure 9 contains the AFD codes for 16:9 coded frames. These are the values that an HD-to-SD down-converter could use to determine automatically the proper picture aspect ratio to use.
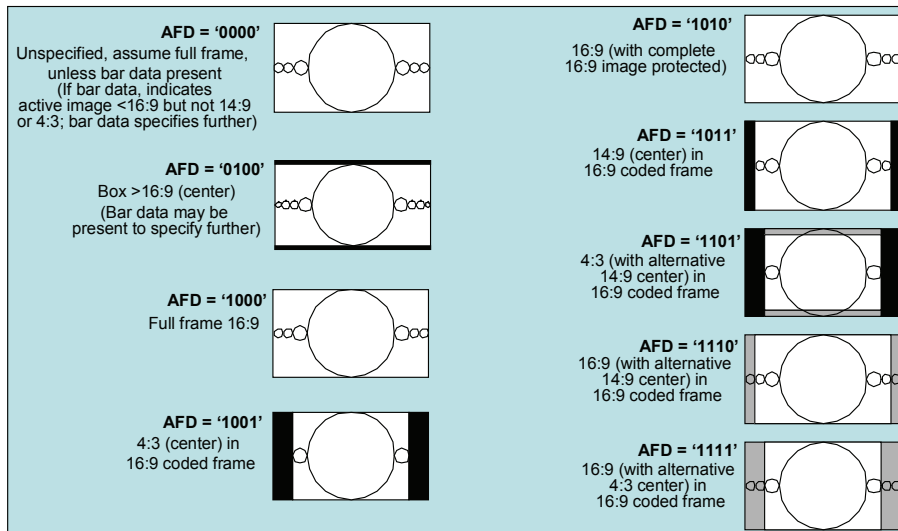

Figure 9. AFD codes for HD scan formats (16:9 code frames). Source: SMPTE 2016-1 [3] (with modifications).

Referring back to the postage stamp example used for Figure 8, if AFD '1001' were signaled along with the picture information, the down-converter would process the HD image as a 4:3 area of interest (protected center), perform the center cut extraction, and down-scale the image to SD. The end result would be a full frame 4:3 SD image (see Figure 10). Neither a postage stamp nor a compromised/destroyed image would occur on a consistent basis (assuming, of course, that at production time the AFD was coded correctly).

Figure 10. Down-conversion using AFD to determine correct picture aspect ratio.

2.2 Bar Data

Bar data is an extension of AFD, for use when the AFD itself does not provide sufficient information to determine the location of the bars (left and/or right in the case of a pillarbox or top and/or bottom in the case of a letterbox). Bar data may be present when AFD codes '0000' or '0100' are used. While not common, applications for bar data include instances when the area of interest is not to be centered in the display raster or when the aspect ratio of the source image is wider than 16:9 (which is often the case for motion pictures, where aspect ratios of 2.39:1, cinemascope, are common). Bar data, if present, contains exactly one pair of bars and specifies the exact location (line number) of the inner extent of the bars (i.e., the last/bottom line of the top bar and the first/top line of the bottom bar). The ATSC-to-NTSC converter/receiver reads the AFD and if bar data is signaled and present, then the active image starts after the first bar (top or left, as applicable) and ends before the second bar (bottom or right, as applicable).

2.3 Color Space Correction

HD and SD content use two different forms of colorimetry – the way that color is represented. HD uses ITU-R Recommendation BT.709 [4], while NTSC uses SMPTE 170M [5]. Another step in the HD to SD down-conversion process is to perform color space correction so that the rendering of the color gamut on the display is correct and reflects what the producer intended. Mathematical models exist for mapping between the two color spaces.

3. Audio Processing

Until 1984, NTSC analog TV sets only supported a single channel of mono audio. In 1984, the FCC adopted the Broadcast Television Systems Committee (BTSC) multi-channel television sound (MTS) standard, which increased the number of supported audio channels and enabled stereo sound fields to be reproduced, all backward compatible with the NTSC-format audio carrier. This is accomplished by encoding the original (main) mono channel as the sum of the Left and Right channels (L+R) such that a legacy NTSC mono set will continue to output the entire audio service. The enhancement channel, which is not processed by a mono TV set, is encoded as the stereo difference of the Left and Right channels (L-R). An analog TV supporting MTS produces the Left channel by summing the main channel (L+R) and the enhancement channel (L-R). The Right channel is produced by taking the difference of the main and enhancement channels ([L+R] – [L-R]). In addition to the main stereo service, MTS defines an additional service, the secondary audio program (SAP). SAP may be used to convey a second audio service in an alternative language or a video description service (for the visually impaired) in mono.

Sound *matrix* methods – where a finite number of discrete audio channels are encoded into fewer channels – were invented such that multi-channel surround sound may be carried on the MTS stereo channels. Dolby Surround is the most common system and encodes four channels of audio (Left, Right, Center, Surround) into a stereo pair (i.e., the MTS main and enhancement channels), using amplitude and phase techniques. When such a surround encoding is done onto a stereo pair, it is referred to as Left total (Lt) and Right total (Rt), as opposed to being normal stereo encoding of Left only (Lo) and Right only (Ro) information. If Lt and Rt are equal and in-phase, the audio is sent to the Center channel (typically for voice sounds on screen). If Lt and Rt are equal but out-of-phase, the audio is sent to the Surround channel (typically for backward noise or special effects). Other combinations of amplitude and phase create a complex sound field. Lastly, summing Lt and Rt produces mono audio (the surround channel information cancels), making it also backward compatible with older analog TV sets that process the main audio channel only. Many analog TV sets available since the 1990s include surround sound processing and five speakers (two surround speakers both for the mono surround channel) and/or provide for outputs to an external audio processor/amplifier.

As part of the ATSC DTV standard, multi-channel surround sound methods were enhanced and the support of multiple audio services was added. The ATSC standardized on the Dolby Digital system (also known as AC-3). Up to six channels of audio – Left, Right, Center, Left Surround, Right Surround, and Low Frequency Effects (LFE) – may be present. Since the LFE channel is limited to frequencies only up to 120Hz, it requires about one-tenth the bandwidth of the other channels, so the six channels are commonly represented as 5.1 channels. These are discrete channels; they are not matrixed. Home theater systems are common today and there is a wide variety of content – from feature length movies to live sporting events – that is produced in 5.1 surround sound.

Analogous to the video down-conversion process, audio processing is required to transform 5.1 surround sound to be compatible with analog TV sets that support only mono or MTS sound. Advances in audio technology enable a single mastering of multi-channel surround sound to deliver the highest quality audio to a wide range of home listening

environments – from specially designed home theater rooms to noisy high-traffic family areas or typical bedrooms. While there are many audio procedures involved, the three main processes required are program loudness adjustment, dynamic range compression, and down-mixing. **Metadata** – information about the audio content – is carried in the ATSC DTV audio bit-stream, along with the audio content. The ATSC-to-NTSC converter/receiver uses the metadata to help in the audio conversion process.

3.1 Program Loudness Adjustment

Simply combining all the audio channels in a multi-channel surround source to create MTS stereo or mono will produce annoyingly loud sound levels. In addition, given the wide variety of source content – some with extremely wide dynamic range (think of a quiet scene of a motion picture with two people having an intimate conversion, followed by an extremely loud explosion) – what is the correct average loudness level to use so that the listener will be comfortable? Both of these issues require that the program loudness be adjusted to a common reference value.

Studies have been done to measure useful objective loudness in order to quantify the correct loudness adjustment [6]. The results of these studies are shown in Figure 11. A **comfort zone** was defined as the range of subjective sound levels around the **comfortable volume point** (defined as 0 dB) in which listeners felt the volume level was acceptable. This is noted in the figure as +2.4 dB to -5.4 dB. When the volume of the program exceeded the upper comfort value, the sound level was considered too loud and caused the listener to lower the volume control of the TV. When the volume of the program exceeded the lower comfort value, the sound level was considered too soft and caused the listener to raise the volume control of the TV. Notice the very narrow range of loudness above the comfortable volume point. This clearly explains why there are so many complaints about loud commercials, as an increase of just 2-3 dB in subjective loudness above the comfort value is enough to be considered unacceptably loud.
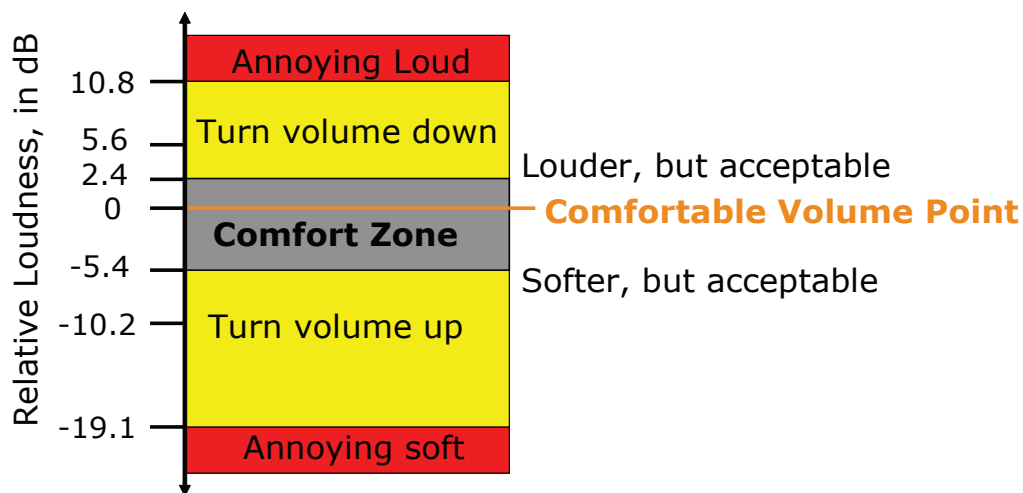

Figure 11. Relative loudness of listening levels, with Comfort Zone identified.

The conclusion of these studies is that listeners most likely adjust TV volume based on the sound levels of normal everyday conversations. As such, Dolby Digital provides a dialogue level normalization metadata parameter called **dialnorm**. Dialnorm represents the long-term A-weighted average level – Leq(A) – of dialogue within a program; that is, the average loudness of dialogue. Leq(A) is measured with a loudness meter. When a Dolby Digital decoder in a receiver processes dialnorm, the average output volume level is normalized to a preset level (i.e., an accepted reference value). If dialnorm is set correctly during production, it will aid in matching volume levels among program sources, whether on the same TV channel or when switching among TV channels.

The accepted reference for average loudness is -31 dBFS Leg(A). Figure 12 shows examples of different types of content and indicates the dynamic range of sound levels, the average dialogue levels, and the signal peaks of each. Note that the feature-length movie has a wide dynamic range where the average dialogue level is relatively low within this range. The rock concert also has a wide dynamic range; however, the average dialogue level is relatively high within this range and, therefore, the end result is a much louder sound level. If the channel were changed from the movie to the rock concert, the volume level on the TV would become unbearably loud.

However, adjusting program loudness by using the dialnorm metadata to normalize the average dialogue level, the same average volume level on the TV would result after the channel change to the rock concert. Figure 13 shows the same audio content examples as Figure 12, but with the average dialogue level normalized. The ATSC-to-NTSC converter/receiver must perform this program loudness adjustment so that the sound being delivered to analog TVs will be acceptable.
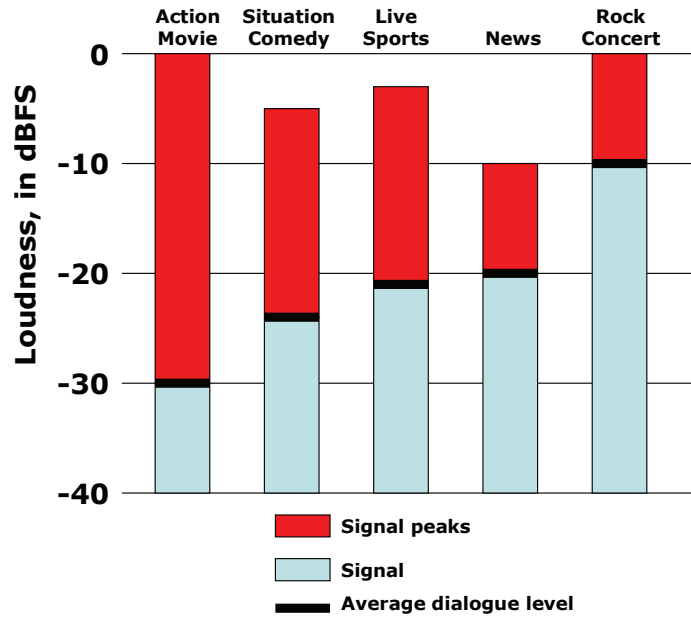

Figure 12. Examples of different types of content, with varying dynamic ranges and average dialogue levels.
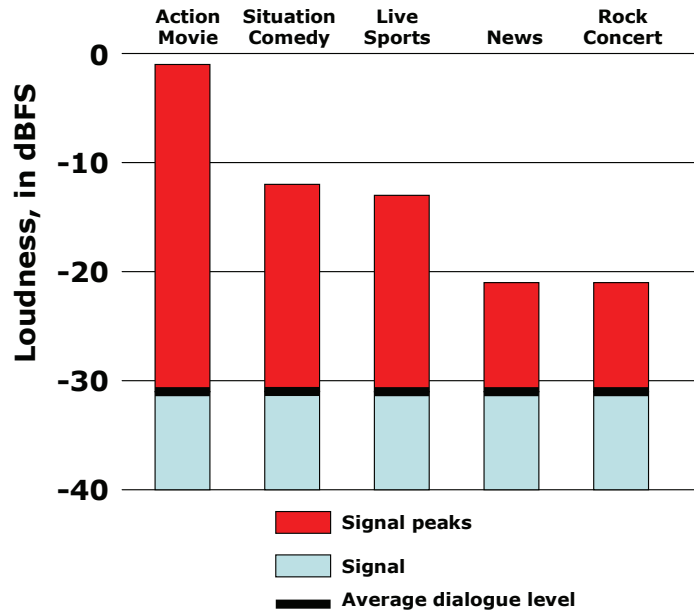

Figure 13. After the dialogue levels have been normalized, the different types of content
have the same average dialogue level, but signal peaks are unaffected.

3.2 Dynamic Range Compression

The Dolby Digital system was designed to enable a "theater" experience. Unfortunately, many home listening environments do not accommodate a wide dynamic range of sound levels. For example, watching the previous movie example on the kitchen analog TV would require the volume control to be turned up during quiet scenes due to typical background/ambient noise and then the explosion scene would likely drive the small mono speaker in the TV into saturation. This not only causes the listener to turn down the excessively loud volume, but also the speaker itself potentially could be damaged. The same program being watched on the large HDTV connected to the home theater system, however, would be enjoyed in its full dynamic range. In most cases, the home listening environment requires that the dynamic range be restricted so that it can be enjoyed in the comfort zone associated with the location in the home and the type of the TV (and associated sound system, if any). As such, dynamic range compression (DRC) must be employed.

Dolby Digital was designed for a single mastering of multi-channel surround sound to deliver the best possible audio for home listening environments, regardless of the audio handling capabilities of the equipment, the ambient noise level in the listening location, or the constraints on the level of practical dynamic range for the listening location. Metadata is provided for two operating modes, **Line Mode** and **RF Mode**. Line Mode typically is used in receivers that have six or two channel line-level outputs. RF Mode typically is used when the audio is RF re-modulated (e.g., Channel 3 or 4 on typical set-top boxes). RF Mode uses heavier dynamic range compression than Line Mode and the peaks are limited. The DRC metadata parameters are **compr** and **dynrng**.

The content producer, or the broadcaster, specifies one of six preset DRC profiles for each mode [7]: Film Light, Film Standard, Music Light, Music Standard, Speech, and None. In the case of the ATSC-to-NTSC converter/receiver, the profile is applied to the appropriate mode selected to produce an acceptable dynamic range for the analog TV.

Each of these profiles applies different characteristics of dynamic range compression, but all follow the general profile shown in Figure 14.
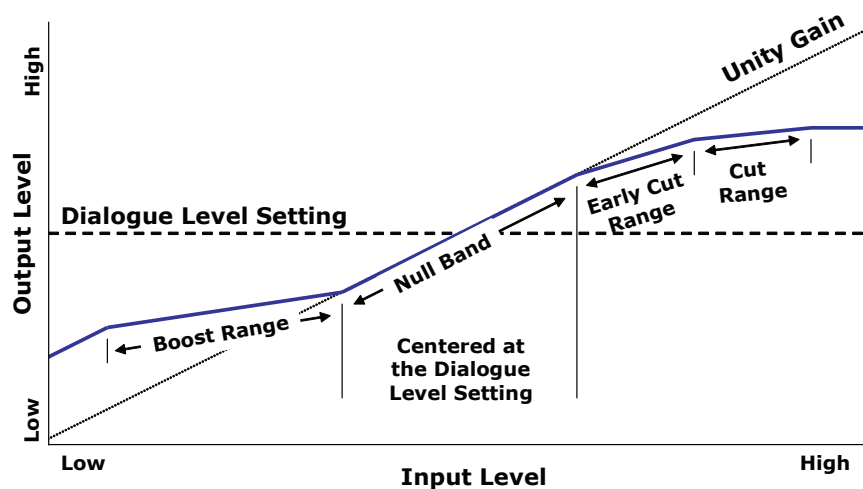

Figure 14. DRC profile general characteristics. Source: Dolby Metadata Guide [7].

Note that the center of the null band is related to the dialogue level setting; hence, the importance of the dialnorm metadata being correct. Automatic overload protection may be applied unnecessarily to prevent clipping if the dialogue level setting is incorrect.

3.3 Down-mixing

Since ATSC audio is multi-channel surround sound and analog TV sets are stereo or mono sound only, audio processing must be done to convert the multi-channel audio to stereo or mono. In the more general case, down-mixing is the process through which multi-channel surround sound is reproduced over fewer speaker channels than for which the program was mastered. The ATSC-to-NTSC converter/receiver must down-mix the Dolby Digital bit-stream to produce the analog stereo output.

The analog stereo output may be either surround compatible – Lt/Rt – for matrix decoding (such as Dolby Surround Pro Logic) or direct stereo down-mix – Lo/Ro – for true stereo separation for speakers or headphones. A mono down-mix is possible as well, and is usually derived from the Lo/Ro down-mix. Typically, down-mixing to stereo outputs uses the Line Mode DRC, while RF Mode DRC is used for mono output.

Down-mixing is not as simple as summing the channels, since unacceptable volume levels may result. Phase relationship is also important for surround-compatible stereo. Overload protection is provided to limit signal peaks to prevent clipping; this feature is not selectable and is always on.

Figure 15 shows the process of down-mixing Dolby Digital to the proper number of speakers present in the receive device, including an analog TV set.
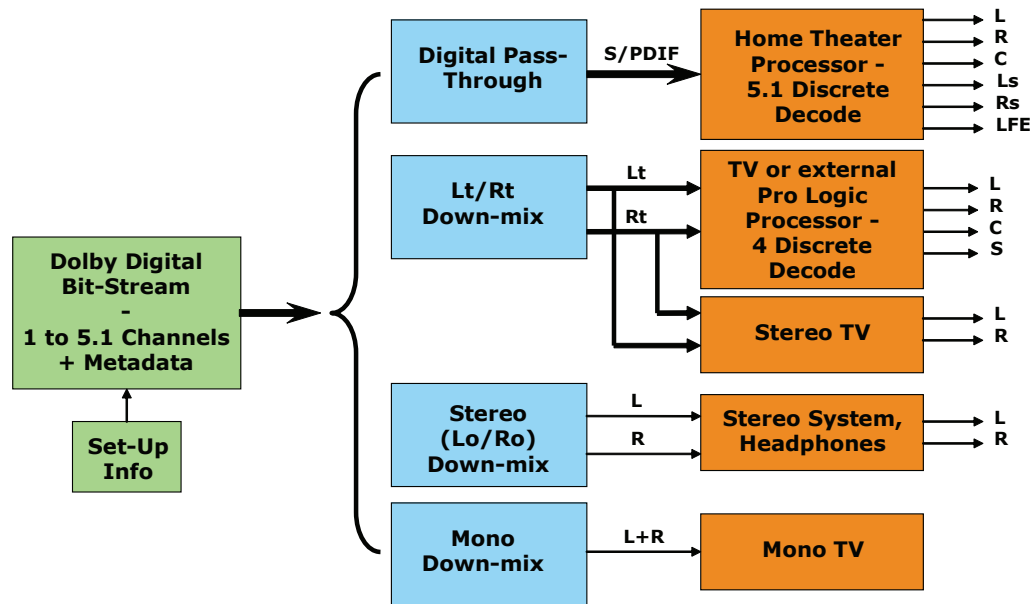

Figure 15. The down-mixing of ATSC surround sound to the proper number of speaker channels.

3.4 Secondary Audio Channel and Video Description Service

The ATSC standard supports multiple audio services, more than just the main/enhancement service and SAP of MTS audio. Moreover, the additional audio services may be multi-channel and not limited to mono like an analog TV set's MTS SAP or video description service. If more than two complete audio services are present, the ATSC-to-NTSC converter/receiver must select which two services are to be used as the primary and SAP. The audio processing of each is independent and, therefore, the DRC and down-mix are done according to whether or not the audio service is the primary service or the SAP or video description service.

4. Closed Captions and Other Ancillary Data

Ancillary data are often overlooked in the discussions of converting content from one format to another. This includes closed captions for the hearing impaired, or for use with the TV mute function, and content advisories (formerly known as program ratings) to give viewers guidance with regard to program content and potential acceptability of the content for children. The use of captions and content advisories are either mandated by the FCC or voluntarily agreed to by the broadcasting and consumer electronics industries.

There are many other types of ancillary data present in an NTSC TV signal. Most are broadcast voluntarily and processed in the TV set or other receive device by industry agreement only. At least one, used by Nielsen Media Research in the preparation of TV program viewer ratings, is considered part of the FCC DTV **must carry** rulings. Two of these are considered important enough to the industries to be considered in the operation of an ATSC-to-NTSC converter/receiver: Nielsen automated measurement of line-ups (AMOL) and TV Guide data.

4.1 Closed Captions

The FCC mandates that TV receivers (whether analog or DTV) with standard displays larger than 13 inches high or widescreen displays larger than 7.8 inches high, and all stand-alone DTV tuners contain built-in decoder circuitry designed to display closed-captioned TV transmissions [8]. The FCC also mandates the inclusion of closed captions in non-exempt programming transmitted by over-the-air broadcasters and other multichannel video programming distributors in a format that will be understandable by the built-in circuitry. The caption format standards are CEA-608 [9] (for NTSC analog TV) and CEA-708 [10] (for ATSC DTV).

Analog TV closed captions are carried on line 21 of the **vertical blanking interval** (VBI) of the NTSC video signal. The VBI is the time difference between the last line of one frame or field of displayed video and the beginning of the next. Originally present to allow time for the electron beam in cathode ray tubes to retrace to the top of the display, the VBI is used as a convenient way to carry ancillary data, synchronized to a particular field/frame, since no active video is displayed during this time. In NTSC video, there are 45 lines in the VBI (about 22 per field) and 480 lines of active video.

CEA-608 defines the data signal on line 21. The data signal may contain up to four closed caption services packets, four text services packets, and one extended data services packet. Only caption packets will be considered in this section. Field 1 and field 2 contain independent data. The primary synchronous caption service – CC1 – is the primary language captioning data that is in sync with a corresponding audio service and it is carried in field 1 only. The secondary synchronous caption service – CC3 – is an alternate captioning data channel, usually for second language captions and, if present, is carried in field 2 only. CC2 and CC4 are special non-synchronous use captions that may be used to augment information carried in the program; they are not often used. Bandwidth is limited to 60 7-bit characters per second per field (about 960 bits per second), with captions having top priority over the other data types. The line 21 data is modulated onto a signal waveform so that it may be transmitted along with the active video waveforms. Within the analog TV set, the caption data is extracted off of the line 21 data signal and, if captions are enabled, rendered into the active video at the display coordinates defined in the caption data and displayed on screen.

For DTV, CEA-708 defines up to 63 caption services, each with more display options (e.g., colors, font sizes, font types) than possible with CEA-608 caption services. Bandwidth, while still constrained, is up to 9,600 bits per second. Unlike CEA-608 captioning that is transported as a waveform in the video VBI, DTV closed captioning (DTVCC) is transported as a logical data channel in the ATSC/DTV bit-stream. Figure 16 shows this pictorially.
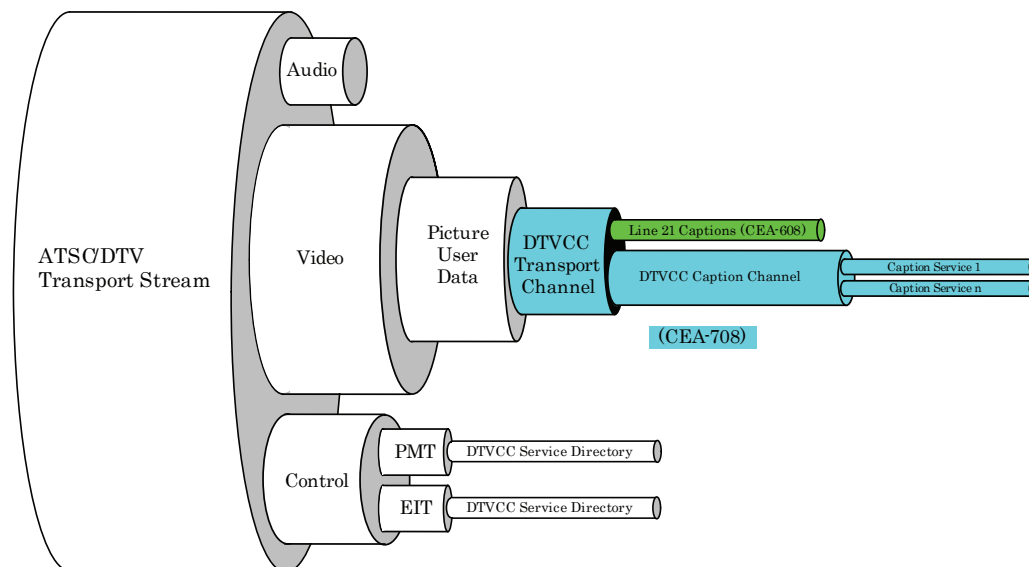

Figure 16. The location of the DTVCC caption data-stream in the ATSC/DTV transport stream. Source: CEA-708 [10] (with modifications)

The **DTVCC Transport Channel** may contain both a CEA-608 caption data-stream (that is, a CEA-608 data-stream embedded in the CEA-708 data-stream, colloquially referred to as "608-in-708") and the DTVCC caption channel. The

DTVCC caption channel may contain up to 63 caption services. The total bandwidth available for all caption data – all DTV caption services and CEA-608 caption services conveyed in the DTVCC Transport Channel – is 9,600 bps, with the CEA-608 captions, if present, being allocated on average about 960 bps.

The DTV Transport Channel and its contents are defined in CEA-708, as is the DTVCC Service Directory, which contains metadata about the caption data-streams. As shown in Figure 16, the DTV Transport Channel is carried as a data-stream in the picture user data area of the video elementary stream for the selected DTV service. Details of the ATSC/DTV transport and video bit-streams are defined in ATSC Standard A/53 Digital Television Standard [11] [12]. The DTVCC Service Directory is carried within DTV in-band control tables, the Program Map Table (PMT) (see A/53) and the Event Information Table (EIT) (see ATSC Standard A/65 Program and System Information Protocol (PSIP) [13]).

The federal regulations governing closed captioning – Title 47, Part 79 – states the "Obligation to pass through captions of already-captioned programs. All video programming distributors shall deliver all programming received from the video programming owner or other origination source containing closed captioning to receiving television households with the original closed captioning data intact in a format that can be recovered and displayed by decoders …". [14]

While not abundantly clear from the legal wording, most experts interpret this to mean that the regulations require the continued use of 608-in-708 caption data-streams in ATSC/DTV signals, if the original source content contains CEA-608 captions. "Native 708" caption service(s) – that is, DTVCC caption service(s) – may be conveyed as well, but not instead of the original CEA-608 captions. The ATSC-to-NTSC converter/receiver, therefore, can extract the embedded CEA-608 data-stream from the incoming DTVCC Transport Channel and reconstruct the line 21 signal waveform on the analog video output. In this manner, analog captions are re-inserted into the analog video.

Some experts believe that the regulations require that both caption formats be provided, 608-in-708 and native 708, so that all TV receivers will be able to decode closed captions.

Lastly, if only native 708 caption data-streams are contained in the incoming ATSC/DTV signal, a conversion to an analog TV signal would cause captions to be lost, or the ATSC-to-NTSC converter/receiver would have to transcode the DTVCC caption service(s) to CEA-608 caption service(s).This would be a complex process and may not result in the required outcome of reproducing equivalent analog caption content, since much of the DTVCC caption format capabilities do not map to the CEA-608 caption format.

4.2 Content Advisories

The FCC mandates that TV receivers with picture screens larger than 13 inches be equipped with features to block the display of TV programming with a common rating, commonly referred to as *V-chip technology*. Broadcasters voluntarily have agreed to broadcast a common program rating system. The rating system, called the *TV Parental Guidelines* (see Figure 17), was established by the National Association of Broadcasters (NAB), the National Cable & Telecommunications Association (NCTA) and the Motion Picture Association of America (MPAA). Similar to analog TV closed captions, the content advisories are transmitted in the VBI of the NTSC video TV signal on line 21 and are displayed on a TV screen in the first 15 seconds of rated programming, and often after a return from a programming break, such as a commercial.

The format and carriage mechanism for analog TV content advisories are defined in CEA-608, the same line 21 services standard used for analog captions. If present, content advisories are carried in line 21, field 2, within an eXtended Data Services (XDS) packet.

The ATSC-to-NTSC converter/receiver applies the same mechanism used to locate CEA-608 captions in the ATSC/DTV signal, extract the data, and then re-generate the line 21 waveform for insertion into the analog video output. In this case, XDS packets containing content advisory information are located in the embedded CEA-608 data-stream in the DTVCC Transport Channel. In fact, to ensure that all line 21 data services in the CEA-608 data-stream are regenerated in the analog TV signal, the ATSC-to-NTSC converter/receiver should extract all line 21 services from the embedded CEA-608 data-stream and regenerate the line 21 waveform in the analog video output.

For DTV content advisories, the mechanism for the carriage of content advisories is very different. The FCC has mandated the use of ATSC A/65 PSIP [13] and CEA-766 *Rating Region Tables and Content Advisory Descriptors for Transport of Content Advisory Information Using PSIP* [15] in its Report and Order of its second periodic review of the analog-to-DTV conversion

[16]. PSIP is used to transport the content advisories. CEA-766 defines the U.S. and Canada rating systems themselves, including the contents of the PSIP Rating Region Tables (RRTs) and the Content Advisory Descriptors (CADs), which provide the specific program rating for each program. The CADs are conveyed in PSIP Event Information Tables (EITs).

TV-Y (All Children -- This program is designed to be appropriate for all children.) Whether animated or live-action, the themes and elements in this program are specifically designed for a very young audience, including children from ages 2-6. This program is not expected to frighten younger children.

TV-Y7 (Directed to Older Children -- This program is designed for children age 7 and above.) It may be more appropriate for children who have acquired the developmental skills needed to distinguish between make-believe and reality. Themes and elements in this program may include mild fantasy or comedic violence, or may frighten children under the age of 7. Therefore, parents may wish to consider the suitability of this program for their very young children. Note: For those programs where fantasy violence may be more intense or more combative than other programs in this category, such programs will be designated TV-Y7-FV.

TV-G (General Audience -- Most parents would find this program suitable for all ages.) Although this rating does not signify a program designed specifically for children, most parents may let younger children watch this program unattended. It contains little or no violence, no strong language and little or no sexual dialogue or situations.

TV-PG (Parental Guidance Suggested -- This program contains material that parents may find unsuitable for younger children.) Many parents may want to watch it with their younger children. The theme itself may call for parental guidance and/or the program contains one or more of the following: moderate violence (V), some sexual situations (S), infrequent coarse language (L), or some suggestive dialogue (D).

TV-14 (Parents Strongly Cautioned -- This program contains some material that many parents would find unsuitable for children under 14 years of age.) Parents are strongly urged to exercise greater care in monitoring this program and are cautioned against letting children under the age of 14 watch unattended. This program contains one or more of the following: intense violence (V), intense sexual situations (S), strong coarse language (L), or intensely suggestive dialogue (D).

TV-MA (Mature Audience Only -- This program is specifically designed to be viewed by adults and therefore may be unsuitable for children under 17.) This program contains one or more of the following: graphic violence (V), explicit sexual activity (S), or crude indecent language (L).

Figure 17. The TV Parental Guidelines rating system. Source: www.fcc.gov.

Unfortunately, an issue occurs with the previously-described mechanism of extracting the XDS content advisory from the DTV Transport Channel. CEA-766-C, section 4, states [text has been paraphrased to increase readability]:

"The source for determining a program's rating using the A/65-defined EIT shall be the CAD contents, including the case of no CAD present in the EIT. Absence of the CAD from the EIT shall mean there is no program rating for the program.

NOTE – This requirement prohibits any use of CEA-608-defined XDS data for blocking when the program is described by the EIT.

NOTE – When a program is not described by an EIT, blocking criteria are outside the scope of CEA-766. Other processing of XDS, including potential blocking by NTSC sets when NTSC outputs are provided from a digital set, is outside the scope of CEA-766. Other blocking systems permitted by FCC rules are outside the scope of CEA-766."

This states clearly that XDS content advisory data shall not be used for blocking a program if a program is described by the EIT, which DTV programs are required to be per the FCC Report and Order. While CEA-766 does not cover the case where a receiving device, such as an ATSC-to-NTSC down-converter/receiver, would pass through the content advisory data to a

receiver further downstream that would perform the blocking, one can easily argue that the extension of this prohibition is appropriate in order to enforce its purpose. That is, since the DTV program is described by the EIT, then the XDS content advisory is not the correct source of the content advisory; it is the PSIP EIT CAD. Simply ignoring the CAD and passing through the XDS content advisory would defeat this.

Granted, CEA-766 also states it is outside its scope if a program is not described by an EIT, with an example given of potential blocking by NTSC sets when NTSC outputs are provided from a DTV set. Could this, therefore, apply to ATSC-to-NTSC converter/receivers? And, as such, would the mechanism previously described for extracting XDS content advisory data from the 608-in-708 data-stream in the DTV Transport Channel and regenerating line 21 XDS content advisory waveform be considered an accepted engineering practice? The ATSC is currently addressing this very issue, with participants from across the industry, including manufacturers, over-the-air broadcasters and cable providers.

Notwithstanding the foregoing industry effort, the ATSC-to-NTSC converter/receiver must provide support for PSIP content advisories in the case where the transmitted DTV transport stream contains PSIP EITs. The ATSC transport stream of the DTV service being converted to analog is parsed to find the proper PSIP EIT, the CAD rating message is extracted, the rating information is converted to XDS content advisory, the line 21 VBI waveform is regenerated and inserted into the analog video. There is a one-to-one mapping of TV Parental Guidelines and MPAA rating system values between CEA-766 CAD rating messages and CEA-608 Content Advisory XDS packet ratings bits. While the process of parsing through the DTV bit-stream to locate the proper PSIP EIT is complicated, it is nonetheless a straight-forward process.
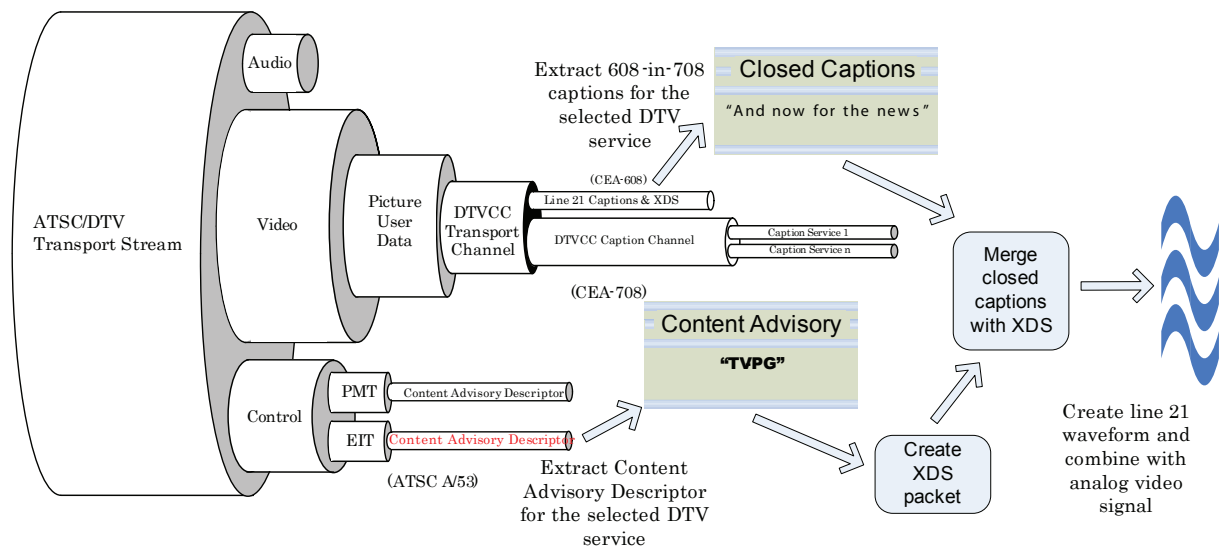

Figure 18. The process (much simplified) of locating the PSIP-based content advisory and creating an XDS content advisory and line 21 VBI waveform.

4.3 Other Ancillary Data

AMOL data (which are used to create the well-known Nielsen TV ratings of program viewers) and TV Guide data (which are used by many analog TV households that do not subscribe to pay TV services) are carried in waveforms on various lines of the VBI. ANSI/SCTE 127 [17] defines a method to carry AMOL-48, AMOL-96 and TVG2X data in the DTV signal. This is shown pictorially in Figure 19. The timing of the VBI data-stream is synchronized with the corresponding video and audio data-streams, as applicable, through the use of common presentation timestamps. The waveforms themselves are defined in CEA-2020 [18]. If present in the incoming ATSC/DTV signal, the ATSC-to-NTSC converter/receiver will parse through the ATSC transport stream, locate the VBI data-stream, extract the AMOL-48, AMOL-96, and/or TVG2X data, regenerate the corresponding VBI waveforms and insert them into the output analog TV signal.
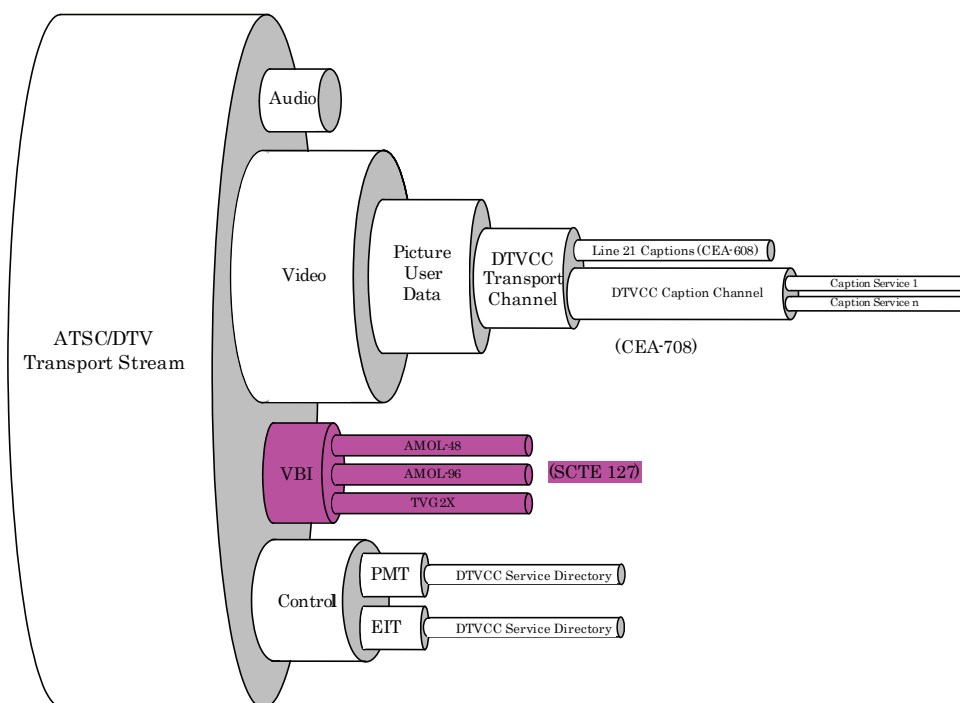
Figure 19. The location of the VBI data-stream in the ATSC/DTV transport stream.

5. Summary

Creating a replacement analog TV signal directly from an ATSC/DTV signal involves many format transformations in order for the video, audio services, closed captions, content advisories, and other ancillary data to be received and processed by an analog TV set. The replacement analog TV signal must take into account:

- HD down-conversion to SD, including picture aspect ratio and color space conversions.
- Audio dynamic range compression, loudness adjustment, multi-channel surround sound down-mix to stereo/mono, and selection of the appropriate services to support SAP and video description.
- Ensuring that the required closed captions and content advisories are present and in the correct format.
- Extracting and passing through Nielsen audience measurement data and TV Guide program data.

ATSC-to-NTSC converter/receivers placed at the front-end of the cable operator's head-end will make it appear to the existing analog plant as if over-the-air analog TV signals were still being broadcast, post the "Great Analog Television Turn-Off" of February 17, 2009. This enables existing cable operators to continue to provide and millions of analog subscribers to continue to enjoy one of their favorite pastimes.

6. RX8320 ATSC Broadcast Receiver

TANDBERG Television's RX8320 ATSC Broadcast Receiver has been developed specifically to help operators re-create analog signals for subscribers after the broadcasters go all-digital in February 2009. The RX8320 offers full translation capability between digital signaling and analog services, and provides all the functions and features of the NTSC-to-ATSC converter/receiver described in the white paper. By offering automatic picture aspect ratio conversion and signaling via AFD and bar data, the RX8320 ensures that widescreen HD video is correctly displayed when down-converted to 4:3 SD video. Full support is provided to ensure that closed captions, Nielsen data, TV Guide data and program rating (V-Chip) services continue to be supported.

For more information on TANDBERG Television's RX8320 ATSC Broadcast Receiver, please see the attached datasheet or visit www.tandbergtv.com.

7. Abbreviations and Acronyms

AFD	Active Format Description
AMOL	Automated Measurement of Line-ups
ATSC	Advanced Television Systems Committee
BTSC	Broadcast Television Systems Committee
C	Center channel
CAD	Content Advisory Descriptor
CEA	Consumer Electronics Association
Dialnorm	Dialogue normalization
DTV	Digital Television
DTVCC	DTV Closed Captioning
DRC	Dynamic Range Compression
EIT	Event Information Table
FCC	Federal Communications Commission
HD	High Definition
L	Left channel
Leq(A)	Long-term A-weighted average level
LFE	Low Frequency Effects
Lo	Left Only
Ls	Left Surround channel
Lt	Left Total mix
MPAA	Motion Picture Association of America
MTS	Multi-channel Television Sound
NAB	National Association of Broadcasters
NTSC	National Television System Committee
NCTA	National Cable & Telecommunications Association
PMT	Program Map Table
R	Right channel
Ro	Right Only
RRT	Rating Region Table
Rs	Right Surround channel
Rt	Right Total mix
PSIP	Program and System Information Protocol
S	Surround channel
SAP	Secondary Audio Programming
SCTE	Society of Cable Telecommunications Engineers
SD	Standard Definition
SMPTE	Society of Motion Picture & Television Engineers
TV	Television
VBI	Vertical Blanking Interval
XDS	eXtended Data Services

8. References

[1] *Media Bureau Staff Report Concerning Over-the-Air Broadcast Television Viewers*, MB Docket No. 04-210, Federal Communications Commission, February 25, 2005.

[2] *Vertical Ancillary Data Mapping of Active Format Description and Bar Data*, SMPTE 2016-3-2007 standard.

[3] *Format for Active Format Description and Bar Data*, SMPTE 2016-1-2007 standard.

[4] *Parameter values for the HDTV standards for production and international programme exchange*, ITU-R Recommendation BT.709-5 (04/02).

[5] *Composite Analog Video Signal – NTSC for Studio Applications*, SMPTE 170M-2004 Standard.

[6] Riedmiller et al, *Intelligent Program Loudness Measurement and Control: What Satisfies Listeners?*, 115th convention of the Audio Engineering Society, October 2003.

[7] *Dolby Metadata Guide*, Issue 3, 2005.

[8] *Report and Order on Closed Captioning Requirements for Digital Television Receivers*, FCC 00-259, Federal Communications Commission, July 21, 2000.

[9] *Line 21 Data Services*, ANSI/CEA-608-E standard, April 2008.

[10] *Digital Television (DTV) Closed Captioning*, CEA-708-C standard, July 2006.

[11] *Digital Television Standard, Part 3: Service Multiplex and Transport Subsystem Characteristics*, ATSC A/53 Part 3 standard, 2007.

[12] *Digital Television Standard, Part 4: Video System Characteristics, with Amendment No. 1*, ATSC A/53 Part 4 standard, 2007.

[13] *Program and System Information Protocol for Terrestrial Broadcast and Cable (Revision C) with Amendment No. 1*, ATSC A/65C standard, January 2006; Amendment 1, May 2006.

[14] p. 750, paragraph (c) of "Closed Captioning and Video Description of Video Programming", 47CFR79.1, Code of Federal Regulations, Title 47, Part 79, October 1, 2007.

[15] *U.S. and Canadian Rating Region Tables (RRT) and Content Advisory Descriptors for Transport of Content Advisory Information Using ATSC Program and System Information Protocol (PSIP)*, ANSI/CEA-766-C standard, April 2008.

[16] *Report and Order on Second Periodic Review of the Commission's Rules and Policies Affecting the Conversion To Digital Television*, FCC 04-192, Federal Communications Commission, August 4, 2004.

[17] *Carriage of Vertical Blanking Interval (VBI) Data in North Americas Digital Television Bitstreams*, ANSI/SCTE 127 2007 standard.

[18] *Other VBI Waveforms¸* ANSI/CEA-2020 standard, December 2007.

TANDBERG television

Part of the **Ericsson** Group