ATSC Standard: A/153 Part 7
AVC and SVC Video System Characteristics

Doc. A/153 Part 7:2012
4 July 2012

Advanced Television Systems Committee
1776 K Street, N.W.
Washington, D.C. 20006
202-872-9160

Revision History

Version	Date
A/153 Part 7:2009, initial version of standard approved	15 October 2009
A/153 Part 7:2012, first revision approved	4 July 2012

Table of Contents

Index of Tables and Figures

ATSC Mobile DTV Standard, Part 7 – AVC and SVC Video System Characteristics (A/153 Part 7:2012)

1. SCOPE

This Part describes a set of video coding constraints on ITU-T Rec. H.264 | ISO/IEC 14496-10 [1] ("AVC") and its Annex G ("SVC") video compression when used in the ATSC Mobile DTV (mobile/handheld, or simply M/H) system [1]. It also defines the RTP packetization for video elementary streams.

1.1 Organization

This document is organized as follows:

- **Section 1** – Outlines the scope of this Part and provides a general introduction.
- **Section 2** – Lists references and applicable documents.
- **Section 3** – Provides a definition of terms, acronyms, and abbreviations for this Part.
- **Section 4** – System overview.
- **Section 5** – Possible video inputs.
- **Section 6** – Source processing before AVC compression.
- **Section 7** – Source coding specification for AVC.
- **Section 8** – Video processing before SVC compression.
- **Section 9** – Source coding specification for SVC.
- **Annex A** – An Example SDP File for Signaling a Two-layer SVC Bitstream in ATSC M/H

2. REFERENCES

All referenced documents are subject to revision. Users of this Standard are cautioned that newer editions might or might not be compatible.

2.1 Normative References

The following documents, in whole or in part, as referenced in this document, contain specific provisions that are to be followed strictly in order to implement a provision of this Part.

[1] ISO: "Information technology -- Coding of audio-visual objects -- Part 10: Advanced Video Coding," Doc. ISO/IEC 14496-10:2010, Geneva, Switzerland.

[2] IEEE/ASTM: "Use of the International Systems of Units (SI): The Modern Metric System," Doc. SI 10-2002, Institute of Electrical and Electronics Engineers, New York, N.Y.

[3] SMPTE: "Format for Active Format Description and Bar Data," Doc. SMPTE 2016-1, Society of Motion Picture and Television Engineers, White Plains, N.Y., 2007.

[4] ITU: "Parameter values for the HDTV Standards for Production and International Programme Exchange," Doc. ITU-R BT.709-5 (2002), International Telecommunications Union, Geneva, Switzerland.

[5] SMPTE: "Standard for Television—Composite Analog Video Signal, NTSC for Studio Applications," Doc. SMPTE 170M (1999), Society of Motion Picture and Television Engineers, White Plains, N.Y.

[6] ATSC: "Video System Characteristics of AVC in the ATSC Digital Television System," Doc. A/72 Part 1:2008, Advanced Television Systems Committee, Washington, D.C., 29 July 2008.

[7] IETF Internet draft (draft-ietf-avt-rtp-rfc3984bis-06.txt), "RTP payload Format for H.264 Video," April 2009. [Editor's note: IETF work in process. RFC number to be assigned.]

[8] IETF Internet draft (draft-ietf-avt-rtp-svc-18.txt), "RTP payload format for SVC video," March 2009. [Editor's note: IETF work in process. RFC number to be assigned.]

[9] IETF: "Signaling media decoding dependency in Session Description Protocol (SDP)," Doc. RFC 5583, Internet Engineering Task Force, Freemont, CA, July 2009.

2.2 Informative References

The following documents contain information that may be helpful in applying this Part.

[10] ATSC: "ATSC Mobile/Handheld Digital Television Standard, Part 1 – Mobile/Handheld Digital Television System," Doc. A/153 Part 1:2011, Advanced Television Systems Committee, Washington, D.C., 1 June 2011.

[11] ATSC: "ATSC Digital Television Standard, Part 2 – RF/Transmission System Characteristics," Doc. A/53 Part 2:2011, Advanced Television Systems Committee, Washington, D.C., 15 December 2011.

[12] ATSC: "ATSC Digital Television Standard, Part 4 – MPEG-2 Video System Characteristics," Doc. A/53 Part 4:2009, Advanced Television Systems Committee, Washington, D.C.,7 August 2009.

[13] SMPTE: "Video Alignment for Compression Coding," Doc. SMPTE RP202, Society of Motion Picture and Television Engineers, White Plains, N.Y., 2007.

[14] CEA: "Digital Television (DTV) Closed Captioning," Doc. CEA-708-D, Consumer Electronics Association, Arlington, VA, August 2008.

[15] ATSC: "ATSC Mobile/Handheld Digital Television Standard, Part 3 – Service Multiplex and Transport Subsystem Characteristics," Doc. A/153 Part 3:2009, Advanced Television Systems Committee, Washington, D.C., 15 October 2009.

[16] IETF: "SDP—Session Description Protocol," Doc. RFC 4566, Internet Engineering Task Force, Freemont, CA.

3. DEFINITION OF TERMS

With respect to definition of terms, abbreviations, and units, the practice of the Institute of Electrical and Electronics Engineers (IEEE) as outlined in the Institute's published standards [2] shall be used. Where an abbreviation is not covered by IEEE practice or industry practice differs from IEEE practice, the abbreviation in question will be described in Section 3.3 of this document.

3.1 Compliance Notation

This section defines compliance terms for use by this document:

shall – This word indicates specific provisions that are to be followed strictly (no deviation is permitted).

shall not – This phrase indicates specific provisions that are absolutely prohibited.

should – This word indicates that a certain course of action is preferred but not necessarily required.

should not – This phrase means a certain possibility or course of action is undesirable but not prohibited.

3.2 Treatment of Syntactic Elements

This document contains symbolic references to syntactic elements used in the audio, video, and transport coding subsystems. These references are typographically distinguished by the use of a different font (e.g., restricted), may contain the underscore character (e.g., sequence_end_code) and may consist of character strings that are not English words (e.g., dynrng).

3.2.1 Reserved Elements

One or more reserved bits, symbols, fields, or ranges of values (i.e., elements) may be present in this document. These are used primarily to enable adding new values to a syntactical structure without altering its syntax or causing a problem with backwards compatibility, but they also can be used for other reasons.

The ATSC default value for reserved bits is '1.' There is no default value for other reserved elements. Use of reserved elements except as defined in ATSC Standards or by an industry standards setting body is not permitted. See individual element semantics for mandatory settings and any additional use constraints. As currently-reserved elements may be assigned values and meanings in future versions of this Standard, receiving devices built to this version are expected to ignore all values appearing in currently-reserved elements to avoid possible future failure to function as intended.

3.3 Acronyms and Abbreviations

The symbols, abbreviations, and mathematical operators used herein are as found in Section 3 of ATSC A/153 Part 1 [10].

4. SYSTEM OVERVIEW

Please see ATSC A/153 Part 1 [10] for an overall description of the M/H system. The ATSC Mobile/Handheld service (M/H) shares the same RF channel as a standard ATSC broadcast service described in ATSC A/53 [11]. M/H is enabled by using a portion of the total available ~19.4 Mbps bandwidth and utilizing delivery over IP transport. The overall ATSC broadcast system including standard (TS Main) and M/H systems is illustrated in Figure 4.1.

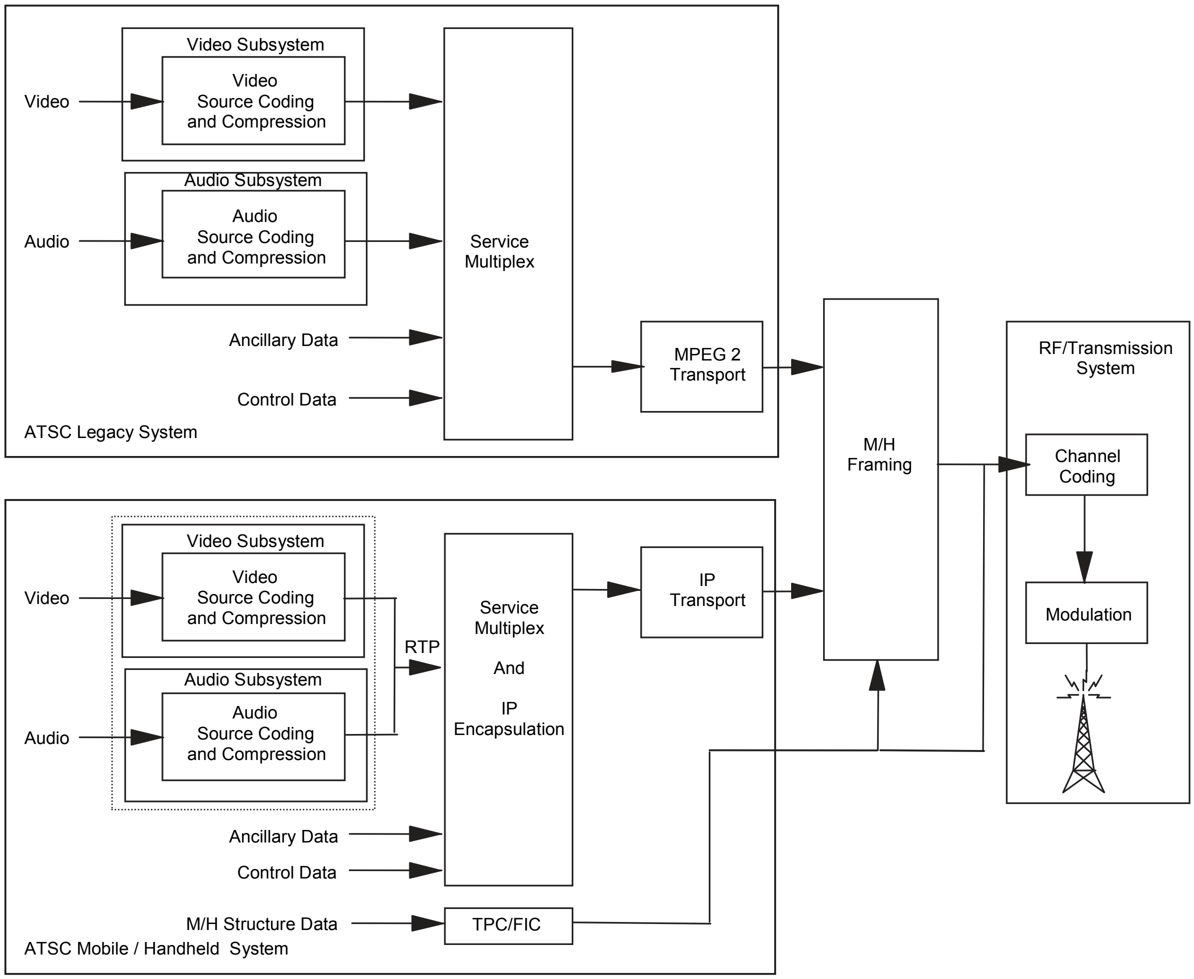

Figure 4.1 ATSC broadcast system with TS Main and M/H services.

This Part relates to the Video Source Coding and Compression block and specifies video coding using MPEG-4 AVC and SVC as described in ISO/IEC 14496 Part 10 [1], with the constraints indicated herein.

5. POSSIBLE VIDEO INPUTS

Please see the A/53 Part 4 Section titled "Possible Video Inputs" [12] for information regarding some television production standards. Production standards supported by this standard include formats with frame rates of 12, 12/1.001 (11.99), 12.5, 15, 15/1.001 (14.98), 24/1.001 (23.976), 24, 25, 30/1.001 (29.97), 30, 50, 60/1.001 (59.94), and 60 Hz. The desired image formats for compression may be derived as indicated in Sections 6 and 8 below.

6. VIDEO PROCESSING BEFORE AVC COMPRESSION

The image formats for AVC compression may be derived from the production video formats as follows.

6.1 1080i Formats

In order to maintain square pixels and simple-ratio scaling factors, of the 1920 pixels per line of video, 24 pixels on the left side of the image and 24 pixels on the right side of the image would need to be cropped[1]. The resulting 1872 pixel by 1080 line image would then be de-interlaced and appropriately re-sampled to 416 pixels by 240 lines prior to compression.

6.2 720p Formats

In order to maintain square pixels and simple-ratio scaling factors, of the 1280 pixels per line of video, 16 pixels on the left side of the image and 16 pixels on the right side of the image would need to be cropped[1]. The resulting 1248 pixel by 720 line image would then be appropriately re-sampled to 416 pixels by 240 lines prior to compression.

6.3 Standard Definition (480i and 480p) Formats with 16:9 Aspect Ratio

For standard definition 16:9 video formats, of the 720 pixels per line of video, 8 pixels on the left side of the image and 8 pixels on the right side of the image should be cropped to produce 704 pixels per line[2]. 480 lines of video should be used for coding (recommended coding ranges are listed in SMPTE RP202 [13]). The resulting 704 pixel by 480 line image is expected to be de-interlaced (if necessary) and appropriately re-sampled to 416 pixels by 240 lines prior to compression.

6.4 Standard Definition (480i) Formats with 4:3 Aspect Ratio

The M/H video system encodes only nominally 16:9 video formats; therefore, 4:3 standard-definition video should be converted to a 16:9 video format before compression. This may be done during up-conversion to a high-definition format or by aspect ratio conversion, remaining in standard-definition. The resulting 16:9 frame (which typically will contain a 4:3 active image with “pillar box” bars on each side) would then be converted to 416 pixels by 240 lines as described in Sections 6.1, 6.2, or 6.3, depending on the video format in use. An alternative arrangement might combine the aspect ratio conversion and re-sampling in a single process.

6.5 Color Space Conversion

Image formats for AVC compression shall use ITU-R BT.709 [4] colorimetry. Production video formats that use another colorimetry shall be transformed into ITU-R BT.709 [4] colorimetry.

6.6 Active Format Description

When the active image area in a 16:9 video signal does not fill the full 16:9 frame, Active Format Description (AFD) and (optionally) Bar Data information in accordance with SMPTE 2016-1 [3] should be present in the source video signal. It is expected that such AFD information, and (optional) Bar Data, will be included in the M/H compressed bit stream and used by an M/H receiver to optimize the display of images that do not fill the coded frame.

Bar Data values are specific to a given video format and, if used, will have to be recalculated when an incoming video signal is converted to the AVC compression format. Formats without

[1] This is 2.5 percent of the active line.

[2] This is approximately 2.2 percent of the active line. Dropping 16 pixels for SD signals is consistent with one implementation of MPEG-2 coding in A/53 [12].

controlling source documents, as specified for compression in this M/H standard, shall use the compressed domain line and pixel numbering.

7. SOURCE CODING SPECIFICATION FOR AVC

This section establishes a specific subset of the AVC video compression standard [1]. The AVC video compression algorithm should conform to either the Constrained Baseline Profile or the Main Profile syntax of AVC video (ISO/IEC 14496-10) [1] appropriate to the constraints and exceptions enumerated below. The allowable parameters shall be bounded by the upper limits specified for the Profiles and Levels enumerated below.

Additionally, AVC bit streams shall meet the constraints and specifications specified in Tables 7.1 through 7.4 and further described in Sections 7.1 through 7.4 of this Part.

AVC bit streams shall utilize both the “Supplemental enhancement information (SEI)” and the “Video usability information (VUI)” syntactic elements defined in Annex D and Annex E of ISO/IEC 14496-10 [1]. Decoder design should be made under the assumption that any legal structure as permitted by ISO/IEC 14496-10 [1] may occur in the broadcast stream even if presently reserved or unused.

7.1 Constraints with Respect to AVC (ISO/IEC 14496-10)

The tables in the following sections list the allowed values for each of the ISO/IEC 14496-10 [1] syntactic elements which are constrained.

7.1.1 Constraints with Respect to AVC Profile

Sequence parameter sets shall be constrained as shown in Section 7.1.3. Picture parameter sets shall be constrained according to Section 7.1.4. In addition, Flexible Macroblock Ordering (FMO) and Arbitrary Slice Order (ASO) shall be disallowed.

7.1.2 AVC Access Point

An Access Point is defined as an access unit in an AVC bit stream at which a decoder can begin decoding video successfully. The access unit must contain one Sequence Parameter Set NAL unit and one Picture Parameter Set NAL unit that are active or being activated when decoding the primary coded picture in this access unit. The access unit must contain an IDR picture or one or more I slices.

7.1.3 Sequence Parameter Set Constraints

For each Access Point, there shall be one Sequence Parameter Set present in the bit stream. Table 7.1 identifies parameters in the Sequence Parameter Set constrained by this specification, with the values that shall be allowed for each.

Table 7.1 Sequence Parameter Set Constraints for AVC

Sequence Parameter Set Syntactic Element	Allowed Value	
	Constrained Baseline	Main
profile_idc	See Table 7.4	See Table 7.4
level_idc	See Table 7.4	See Table 7.4
constraint_set0_flag	1	0
constraint_set1_flag	1	1
constraint_set2_flag	1	0
constraint_set3_flag	0	0
PicWidth InMbs	See Table 7.4	See Table 7.4
PicHeight InMbs	See Table 7.4	See Table 7.4
aspect_ratio_idc	See Table 7.4	See Table 7.4
num_units_ in_tick	See Table 7.3	See Table 7.3
time_ scale	See Table 7.3	See Table 7.3

Note: These are the values of the parameters for signaling the use of this video codec constraint set for AVC.

The constraint_set0_flag shall be set to '1' for Constrained Baseline Profile and to '0' for Main Profile. The constraint_set1_flag shall be set to '1'. The constraint_set2_flag shall be set to '1' for Constrained Baseline Profile and to '0' for Main Profile. The constraint_set3_flag shall be set to '0'. The time interval between two changes in pairs of pic_width_in_mbs_minus1 and pic_height_in_map_units_minus1 shall be greater than or equal to one second.

7.1.4 Picture Parameter Set Constraints

More than one Picture Parameter Set can be present in the bit stream between two Access Points. Between two Access Points, the content of a Picture Parameter Set with a particular pic_parameter_set_id shall not change. (If more than one Picture Parameter Set is present in the bit stream and these Picture Parameter Sets are different from each other, then each picture parameter set shall have a different pic_parameter_set_id.). Table 7.2 identifies parameters in the Picture Parameter Set of a bit stream constrained by this Part, with the values that shall be allowed for each.

Table 7.2 Picture Parameter Set Constraints for AVC

Picture Parameter Set Syntactic Element	Allowed Value
num_slice_group_minus1	0
redundant_pic_cnt_present_flag	0

7.1.5 Video Usability Information (VUI) Parameter Constraints

The following parameters in the Video Usability Information (VUI) part of a bit stream that shall be constrained are:

- video_format – shall only take the value of '000'
- low_delay_hrd_flag – shall only take the value of '0'.

The decoder for this video shall support the use of the VUI's following syntax elements:

- Aspect Ratio Information (aspect_ratio_idc)
- Color Parameter Information (colour_primaries, transfer_characteristics, and matrix_coefficients)

- Chrominance Information (chroma_sample_loc_type_top_field and chroma_sample_loc_type_bottom_field)
- Timing information (time_scale, num_units_in_tick, low_delay_hrd_flag, timing_info_present_flag, and fixed_frame_rate_flag).

The values for color_primaries, transfer_characteristics, and matrix_coefficients shall be explicitly indicated in the vui_parameters( ). The values for color_primaries, transfer_characteristics, and matrix_coefficients defined for ITU-R BT.709 [4] shall be used.

The values for time_scale, num_units_in_tick, and fixed_frame_rate_flag shall be explicitly indicated in the vui_parameters( ). Table 7.3 indicates the defined frame rates, and the values for num_units_in_tick and time_ scale that shall be used to signal them.

Table 7.3 Frame Rate VUI Parameter Constraints for AVC

Frame Rate	num_units_in_tick	time_ scale
11.99	1001	24000
12 Hz	1	24
12.5 Hz	1	25
14.98 Hz	1001	30000
15 Hz	1	30
23.98 Hz	1001	48000
24 Hz	1	48
25 Hz	1	50
29.97 Hz	1001	60000
30 Hz	1	60

7.2 Compression Format Constraints

This Part supports compression formats with horizontal sizes and vertical sizes as defined in Table 7.4 with frame rates as defined in Table 7.3. The aspect_ratio_idc shall equal 1 (square samples). The display aspect ratio shall be nominally 16:9.

Table 7.4 lists the compression format details.

Table 7.4 Compression Format Constraints for AVC

Profile	Vertical Size	Horizontal Size	PicWidth in Mbs	PicHeight in Mbs	aspect_ ratio_idc	profile_ idc	level_ idc	Display aspect ratio	Progressive/ interlaced
Constrained Baseline	240	416	26	15	1	66	13	16:9	P
Main	See Note 1	See Note 1	1 to 80	1 to 45	1	77	31	16:9	P
Note 1: Vertical and Horizontal sizes shall be within the constraints of ISO/IEC 14496-10 [1], where the maximum number of macroblocks is 80 x 45, in a ratio as close to 16:9 as possible.									

Examples of transmitted formats that are within the constraints of this section are:

576V x1024 H

480V x 832H

368V x 640H

288V x 512H

192V x 336H

7.3 Low Delay and Still Picture Modes

Low delay and still picture requirements shall be according to ATSC A/72 Part 1 Section 6.3 [6].

7.4 Bit Stream Specifications for Closed Captioning, AFD, and Bar Data

Closed Captioning, Active Format Description and Bar Data shall be carried in the SEI_RBSP and VUI sections of the video syntax as described in ISO/IEC 14496-10 [1]. For Closed Captioning, the usage shall be according to ATSC A/72 Part 1 Section 6.4. [6], except that variable bit rates, not to exceed 9600 bits per second, shall be permitted for the closed caption payload (that is, packing bytes need not be used, and when captions are not present there is no bandwidth allocation).

> Note: This is an intentional difference from the current version of CEA-708 [14], which requires allocation of 9600 bits per second for the closed caption payload data for all DTV-system bit streams.

For AFD and Bar Data, the usage shall be according to ATSC A/72 part 1 Section 6.4 [6], except that compressed domain line and pixel numbering shall be used for Bar Data (see Section 6.6 for video preprocessing requirements).

7.5 RTP Packetization

The AVC video elementary stream shall be carried in NAL units according to ISO/IEC 14496-10 [1], and transported with an RTP payload format according to [7], with the following additional constraints.

- Interleaved packet mode shall not be used.
- An NAL unit of small size should be encapsulated in an aggregation packet together with one or more other NAL units, whenever it is allowed by [7]. As a consequence, the use of Single-Time Aggregation Packet (STAP-A) is encouraged over the use of Single NAL Unit Packet. NAL units which will not fit into a single MTU should use Fragmented Single NAL Unit Packets (FU-A).

> Note: These requirements do not take into account packet re-ordering that may or may not occur outside the ATSC-M/H system.

In addition, for RTP packets that carry AVC video elementary stream, the payload_type field in the RTP header shall have value 35.

> Note: The signaling of the AVC configuration is defined in A/153 Part 3 [15].

8. VIDEO PROCESSING BEFORE SVC COMPRESSION

The image formats for the SVC base layer (see Section 9) compression may be created from the production video formats as indicated in Section 6.

The image formats for the SVC bit stream (see Section 9) compression may be created from the production video formats as follows.

8.1 1080i Formats

In order to maintain square pixels and simple-ratio scaling factors, of the 1920 pixels per line of video, 24 pixels on the left side of the image and 24 pixels on the right side of the image would need to be cropped[3]. The resulting 1872 pixel by 1080 line image would then be de-interlaced and appropriately re-sampled to 832 pixels by 480 lines, or to 624 pixels by 360 lines, prior to compression.

8.2 720p Formats

In order to maintain square pixels and simple-ratio scaling factors, of the 1280 pixels per line of video, 16 pixels on the left side of the image and 16 pixels on the right side of the image would need to be cropped[3]. The resulting 1248 pixel by 720 line image would then be appropriately re-sampled to 832 pixels by 480 lines, or to 624 pixels by 360 lines, prior to compression.

8.3 Standard Definition (480i and 480p) Formats with 16:9 Aspect Ratio

For standard-definition 16:9 video formats, of the 720 pixels per line of video, 8 pixels on the left side of the image and 8 pixels on the right side of the image should be cropped to produce 704 pixels per line[4]. 480 lines of video should be used for coding (recommended coding ranges are listed in SMPTE RP202 [13]). The resulting 704 pixel by 480 line image is expected to be de-interlaced (if necessary) and up-sampled to 832 pixels by 480 lines, or de-interlaced (if necessary) and appropriately re-sampled to 624 pixels by 360 lines, prior to compression.

8.4 Standard Definition (480i) Formats with 4:3 Aspect Ratio

The M/H video system encodes only nominally 16:9 video formats; therefore, 4:3 standard-definition video should be converted to a 16:9 video format before compression. This may be done during up-conversion to a high-definition format or by aspect ratio conversion, remaining in standard-definition. The resulting 16:9 frame (which typically will contain a 4:3 active image with “pillar box” bars on each side) would then be converted to 832 pixels by 480 lines, or 624 pixels by 360 lines as described in Sections 8.1, 8.2, or 8.3, depending on the video format in use. An alternative arrangement might combine the aspect ratio conversion and re-sampling in a single process.

8.5 Color Space Conversion

Image formats for SVC compression shall use ITU-R BT.709 [4] colorimetry. Production video formats that use another colorimetry shall be transformed into ITU-R BT.709 [4] colorimetry.

8.6 Active Format Description

When the active image area in a 16:9 video signal does not fill the full 16:9 frame, Active Format Description (AFD) and (optionally) Bar Data information in accordance with SMPTE 2016-1 [3] should be present in the source video signal. It is expected that such AFD information, and (optional) bar data, will be included in the M/H compressed bit stream for the SVC base layer only and not in the SVC enhancement layer (see Section 9.4). M/H receivers are

---

[3] This is 2.5 percent of the active line.

[4] This is approximately 2.2 percent of the active line. Dropping 16 pixels for SD signals is consistent with one implementation of MPEG-2 coding in A/53 [12].

expected to optimize the display of images carried by the SVC bit stream by using AFD and Bar Data that may be present in the M/H bit stream for the SVC base layer.[5]

Bar Data values are specific to a given video format and, if used, will have to be recalculated when an incoming video signal is converted to the SVC compression format. Although the video image spatial processing before SVC compression is based on the format for the SVC enhancement layer, it should be noted that the recalculated Bar Data line and pixel numbers will have to be based on the SVC base layer format. Formats without controlling source documents, as specified for compression in this M/H standard, shall use the compressed domain line and pixel numbering.

9. SOURCE CODING SPECIFICATION FOR SVC

This section establishes a specific subset of the SVC Annex of the video compression standard [1]. SVC is an optional capability for M/H encoding and M/H decoding devices but, when implemented, shall conform to the requirements specified herein.

The SVC bit stream comprises two layers: the term "SVC base layer" is used to designate the layer that has dependency_id and quality_id values both equal to 0, and the term "SVC enhancement layer" to designate the layer that has dependency_id value higher than 0.

With the exceptions stated in the following paragraph, the SVC video compression algorithm shall conform to the Scalable Baseline Profile of SVC video, Annex G of ISO/IEC 14496-10 [1]. The allowable parameters for the SVC bit stream shall be bounded by the upper limits specified for Scalable Baseline Profile at Level 3.1 defined in Annex G of ISO/IEC 14496-10 [1].

The SVC base layer shall meet the constraints defined in Section 7 for AVC source coding. Additionally, the SVC bit stream shall meet the constraints and specifications specified in Table 9.1 and Table 9.2 and as further described in Sections 9.1 through 9.5 of this Part.

SVC bit streams shall utilize the "Supplemental Enhancement Information (SEI)" defined in Annexes D and G of ISO/IEC 14496-10 [1], and the "Video Usability Information (VUI)" defined in Annex E and Annex G of ISO/IEC 14496-10 [1]. Decoder design should be made under the assumption that any legal structure as permitted by Annex G of ISO/IEC 14496-10 [1] may occur in the broadcast stream even if presently reserved or unused.

9.1 Constraints with Respect to SVC (ISO/IEC 14496-10, Annex G)

9.1.1 Constraints with Respect to SVC Baseline Profile

Each SVC bit stream shall contain not more than two distinct values of dependency_id. For the SVC enhancement layer, each SVC VCL NAL unit (NAL unit type 20) shall have the values of dependency_id and quality_id equal to 1 and 0, respectively.

[5] Since the active format of the enhancement layer is always identical to that of the base layer, a receiver processing the enhancement layer SVC picture in the presence of a base layer AFD code is expected to handle it in the same as a receiver processing the base layer SVC picture. The receiver is expected to recalculate the Bar Data line and pixel numbers (if used) relating to the enhancement layer video format, based on the SVC base layer Bar Data (the inverse process of that mentioned for pre-compression processing in the second paragraph of 8.6).

9.1.2 SVC Access Point

In SVC context, an Access Point is associated with a target dependency representation for output. Specifically, at an SVC Access Point for a particular value of dependency_id, the decoder can begin decoding all the dependency representations that have dependency_id less than or equal to the target dependency representation.

The access unit that corresponds to an SVC Access Point for a particular value of dependency_id shall contain a dependency representation that has the particular value of dependency_id and idr_flag is equal to 1. At each Access Point for the SVC bit stream, the access unit shall contain all SVC Sequence Parameter Set NAL units (NAL unit type 7 or 15) and all Picture Parameter Set NAL units (NAL unit type 8) that are referenced in the VCL NAL units of the access unit. The access unit shall not contain any Sequence Parameter Set NAL unit (NAL unit type 7) that is not referenced in the VCL NAL units of the access unit. In addition, if both are present in the access unit, the Sequence Parameter Set NAL unit associated with the SVC base layer shall precede the Subset Sequence Parameter Set NAL unit associated with the SVC enhancement layer.

9.1.3 Sequence Parameter Set Constraints

For each Access Point in the SVC bit stream, there shall be one Subset Sequence Parameter Set present in the bit stream. Table 9.1 identifies parameters in the SVC enhancement layer Subset Sequence Parameter Set, with the values that shall be allowed for each.

Table 9.1 Subset Sequence Parameter Set Constraints for SVC

Subset Sequence Parameter Set Syntactic Element	Allowed Value
profile_idc	83
level_idc	See Table 9.2
PicWidth InMbs	See Table 9.2
PicHeight InMbs	See Table 9.2
aspect_ratio_idc	See Table 9.2
num_units_in_tick	See Table 7.3
time_scale	See Table 7.3

Note: These are the values of the parameters for signaling the use of this video codec constraint for the SVC enhancement layer.

The time interval between two changes in pairs of pic_width_in_mbs_minus1 and pic_height_in_map_units_minus1 shall be greater than or equal to 1 second.

9.1.4 Picture Parameter Set Constraints

More than one Picture Parameter Set can be present in the SVC bit stream between two Access Points. Between two Access Points, the content of a Picture Parameter Set with a particular pic_parameter_set_id shall not change. (If more than one Picture Parameter Set is present in the bit stream and these picture parameter sets are different from each other, then each picture parameter set shall have a different pic_parameter_set_id).

In addition, Table 9.2 identifies the parameters in the Picture Parameter Set of the bit stream of the SVC base layer that are constrained by this Part, with the values that shall be allowed for each.

Table 9.2 Picture Parameter Set Constraints for SVC Base Layer

Picture Parameter Set Syntactic Element	Allowed Value
constrained_intra_pred_flag	1
num_slice_group_minus1	0
redundant_pic_cnt_present_flag	0

9.1.5 SVC Video Usability Information (VUI) Parameter Extension Constraints

The constraints for SVC VUI parameter extension shall follow the constraints specified in Section 7.1.5. The values for time_scale, num_units_in_tick, and fixed_frame_rate_flag as specified in Table 7.3 shall be extended to include the corresponding values for 50 Hz, 59.94 Hz, and 60 Hz frame rates specified by Table 6.2 of A/72 Part 1 [6]. In addition, the value of vui_ext_num_entries_minus1 shall be equal to or greater than 1.

9.2 Compression Format Constraints

The compression formats for the SVC base layer shall be as defined in Section 7.2

Table 9.3 lists the compression format details for the SVC enhancement layer.

Table 9.3 Compression Format Constraints for SVC

vertical size	horizontal size	PicWidth InMbs	PicHeight InMbs	aspect_ ratio_idc	profile_ idc	level_ idc	display aspect ratio	allowed frame rate	Progressive Interlaced
360[1]	624	39	23	1	83	31	16:9[2]	*See legend*	P
480	832	52	30	1	83	31	16:9[2]	*See legend*	P

Legend:
Supported frame rates: 11.99 Hz, 12 Hz, 12.5 Hz, 14.98 Hz, 15 Hz, 23.98 Hz, 24 Hz, 25 Hz, 29.97 Hz, 30 Hz, 50 Hz, 59.94 Hz, 60 Hz

Notes:
1 368 lines are actually coded in order to satisfy the AVC/SVC requirement that the coded vertical size be a multiple of 16. The bottom 8 lines are black per MPEG rules.
2 Actually 15.6:9, based on vertical and horizontal sizes and square pixels.

9.3 Low Delay and Still Picture Modes

Refer to Section 7.3.

9.4 Bit Stream Specifications for Closed Captioning, AFD, and Bar Data

Closed Captioning, AFD, and Bar Data shall not be carried in the bit stream of the SVC enhancement layer. Receivers decoding bit stream of the SVC enhancement layer are expected to use the information that may be present in the bit stream for the SVC base layer. (See Section 7.4.)

9.5 RTP Packetization

The SVC video elementary stream may be delivered in one or two separate RTP sessions for transport, depending on application requirements.

In both cases, NAL units in the bit stream shall be packetized into RTP packets according to [8], with the following additional constraints:

- Interleaved mode shall not be used.
- Each prefix NAL unit (NAL unit type 14) should be included in the same aggregation packet when an aggregation packet is used for its associated NAL unit (NAL unit type 1 or 5), with the exception of the following cases:
  - The formed aggregation packet violates session MTU constraint.
  - Fragmentation units are used for the associated NAL unit.

In addition, the RTP session(s) to deliver the SVC video elementary stream shall meet the constraints and specifications further specified in Sections 9.5.1 and 9.5.2 of this document.

Note: The signaling of the SVC configuration is defined in A/153 Part 3 [15].

9.5.1 SVC Transport in Two RTP Sessions

In this case, the SVC video elementary stream shall be separated into one SVC base layer elementary stream and one SVC enhancement layer elementary stream based on NAL unit types.

The SVC base layer elementary stream shall contain NAL units associated with the SVC base layer, including NAL unit types of 1, 5, 7, 8 and 14. The SVC enhancement layer elementary stream shall contain NAL units associated with the SVC enhancement layer, including NAL unit types of 15 and 20. If present in the elementary stream, an NAL unit with type of 6, 9, 10, 11, or 13 shall be included in the SVC base layer. A NAL unit with type of 12 shall be included in the layer which its immediately preceding VCL NAL unit belongs to.

The resulting SVC base and enhancement layer elementary streams shall be delivered in two RTP sessions separately, with the following constraints:

- Both sessions shall obey the constraints for Multi-Session Transmission (MST) specified in [8].
- The SVC base layer session shall follow the RTP packetization constraints specified in Section 7.5.
- Single NAL unit packetization mode shall not be allowed for the enhancement layer.
- Both sessions shall have the same synchronization source (SSRC) identifier. The RTP session corresponding to the SVC base layer shall be used for SSRC identifier allocation and collision resolution.
- The RTP timestamps for each session shall be synchronized; i.e., both sessions should choose the same random initial value for timestamps. Furthermore, the timestamps for both sessions shall be derived from the same clock instant, so that the RTP packets from different sessions have the same timestamp if they belong to the same video frame.
- Each RTP session shall have an associated RTCP session.
- The requirements for signaling decoding dependency between the two sessions shall be as defined in [8] and [9].

In addition, the RTP session that delivers SVC base layer elementary streams shall conform to the RTP packetization constraints specified in Section 7.5. As a result, the payload_type field in the RTP header of each RTP packet within the session that delivers SVC base layer elementary streams shall have value 35. Furthermore, for the RTP session that delivers SVC enhancement layer elementary streams, the payload_type field in the RTP header of each RTP packet within the session shall have value 36.

Note: ATSC-M/H Signaling is defined in A/153 Part 3 [15].

9.5.2 SVC Transport in Single RTP Session

In this case, the RTP session that contains the SVC video elementary stream shall have an associated RTCP session. That session shall obey the constraints for Single-Session Transmission (SST) specified in [8]. Furthermore, the use of Single-Time Aggregation Packet (STAP-A) is encouraged over the use of Single NAL Unit Packet.

In addition, for the RTP packets that carry NAL units corresponding to the SVC base layer as specified in Section 9.5.1, the payload_type field in the RTP header shall have value 35. Furthermore, for the RTP packets that carry NAL units corresponding to the SVC enhancement layer as specified in Section 9.5.1, the payload_type field in the RTP header shall have value 36.

Annex A: Relationship between MH_Component_Data Descriptor and SDP

ATSC M/H transmits SDP messages according to RFC 4566 [16] for announcement of services. For signaling of video codec capabilities, however, the MH_component_descriptor() with the MH_component_data() structure (for Component Type 35) as defined in Sections 7.8.1.1 and 7.8.1.2 of A/153 Part 3 [15] is used. The MH_component_data() structure and the SDP messages carry many of the same parameters. It is strongly recommended to use the MH_component_descriptor() with the MH_component_data() structure for initialization of the video decoder, because the MH_component_descriptor() is defined to take precedence over the SDP message.

The following section explains the elements in an example SDP message to help clarify how the video signaling of a two-layer SVC bitstream in ATSC M/H is related to SDP. Consider the following SDP message:

```
1   c = IN IP4 192.0.2.1 / 127
2   a = group:DDP L1 L2

3   m = video 40000/2 RTP / AVP 35
4   a  = rtpmap:35   H264 / 90000
5   a  = fmtp:35 profile-level-id = 42e00d; sprop-parameter-sets = {sps0}, {pps0}; packetization-mode=0;
6   a  = mid : L1

7   m = video 40002/2 RTP/AVP 98
8   a = rtpmap:36   H264-SVC/90000
9   a = fmtp:36  profile-level-id = 53001f; sprop-parameter-sets = {ssps1}, {pps1}; packetization-mode=1;
10  a = mid : L2
11  a = depend : 36   lay L1 : 35
```

Within this SDP message,

1) Lines 3 – 6 and 7 – 11 describe the session information for the SVC base layer and the SVC enhancement layer respectively, and Lines 1 – 2 are shared by both layers. In addition, for an AVC-only receiver, it shall only use the relevant information from the common and the SVC base layer sections, in order to receive and decode the AVC bitstream, which is a subset of the SVC base layer.
2) Line 2 indicates that the decoding dependency group (DDP) consists of two streams; i.e., L1 and L2.
3) Lines 3 – 6 describe the base layer part of the SVC bitstream. The bitstream is coded in Constrained Baseline Profile at Level 1.3, and packetized in RTP Single-NAL Unit mode by setting packetization-mode as 0. In Line 5, {sps0} and {pps0} represent the Base64-encoded SPS and PPS NAL units necessary for decoding the SVC base layer. Furthermore, the bitstream is identified by its media description (mid) as L1 in Line 6.
4) Lines 3 – 5 describe the AVC bitstream. The bitstream is coded in Constrained Baseline Profile at Level 1.3, and packetized in RTP Single-NAL Unit mode by setting packetization-mode as 0 ([7], Section 5.4). In Line 5, {sps0} and {pps0} represent the Base64-encoded SPS and PPS NAL units necessary for decoding the AVC bitstream.
5) Lines 7 – 11 describe the enhancement layer part of the SVC bitstream. The bitstream is coded in Scalable Baseline Profile at Level 3.1, and packetized in RTP non-interleaved mode

by setting packetization-mode as 1 ([7], Section 5.4). In Line 9, {ssps1} and {pps1} represent the Base64-encoded SSPS and PPS NAL units necessary for decoding the SVC enhancement layer bitstream. The bitstream is identified by its media description as L2 in Line 10. Finally, the “depend” attribute in Line 11 indicates that the bitstream is dependent on the base layer bitstream L1.

6) In the SDP file, yellow-colored fields are mandatory for the SVC base layer, green-colored fields are mandatory for the SVC enhancement layer, and grey-colored fields are optional but recommended. In particular, SPS, SSPS and PPS NAL units are optionally encoded and given by the sprop-parameter-sets field in Lines 5 and 9. These NAL units not only enable the decoder to correctly decode the received SVC bitstream, but also contain important information for receiver buffer model conformance.