CS/TSG-658r1
9 May 2006

Candidate Standard: Amendment to ATSC Digital Television Standard, Doc. A/53D, to define Video System Characteristics of VC-1 as new Annex F

Advanced Television Systems Committee
1750 K Street, N.W.
Suite 1200
Washington, D.C. 20006
www.atsc.org

The Advanced Television Systems Committee, Inc., is an international, non-profit organization developing voluntary standards for digital television. The ATSC member organizations represent the broadcast, broadcast equipment, motion picture, consumer electronics, computer, cable, satellite, and semiconductor industries.

Specifically, ATSC is working to coordinate television standards among different communications media focusing on digital television, interactive systems, and broadband multimedia communications. ATSC is also developing digital television implementation strategies and presenting educational seminars on the ATSC standards.

ATSC was formed in 1982 by the member organizations of the Joint Committee on InterSociety Coordination (JCIC): the Electronic Industries Association (EIA), the Institute of Electrical and Electronic Engineers (IEEE), the National Association of Broadcasters (NAB), the National Cable Television Association (NCTA), and the Society of Motion Picture and Television Engineers (SMPTE). Currently, there are approximately 140 members representing the broadcast, broadcast equipment, motion picture, consumer electronics, computer, cable, satellite, and semiconductor industries.

ATSC Digital TV Standards include digital high definition television (HDTV), standard definition television (SDTV), data broadcasting, multichannel surround-sound audio, and satellite direct-to-home broadcasting.

About the Candidate Standard

This specification is being put forth as a Candidate Standard by the Specialist Group on Audio and Video Coding (TSG/S6). ATSC members and non-members are encouraged to review and implement this specification and return comments to cs_amend_editor@atsc.org. ATSC Members can also send comments directly to the TSG/S6 Specialist Group. The ATSC believes this specification is stable. It is expected to progress to Proposed Standard within a period of time ending 31 December 2007.

Insert new Annex F of A/53D as follows:

A/53D Annex F:
VC-1 Video System Characteristics (Normative)

1. SCOPE

This Annex describes the characteristics and constraints to the VC-1 video subsystem of the Digital Television Standard. Readers should consult [F7] for the bit stream structure of the coded signal, and should consult the other sections of this document for the integration into the ATSC infrastructure.

2. REFERENCES

2.1 Normative References

The following documents contain provisions which, through reference in this text, constitute provisions of this standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreement based on this standard are encouraged to investigate the possibility of applying the most recent editions of the documents listed below.

[F1] ISO/IEC 13818-1:2000 (E), International Standard, Information technology – Generic coding of moving pictures and associated audio information: Systems.

[F2] SMPTE 125M (1995): Standard for Television—Component Video Signal 4:2:2, Bit-Parallel Digital Interface, Society of Motion Picture and Television Engineers, White Plains, N.Y.

[F3] SMPTE 170M (2004): Standard for Television—Composite Analog Video Signal, NTSC for Studio Applications, Society of Motion Picture and Television Engineers, White Plains, N.Y.

[F4] SMPTE 274M (2005): Standard for Television—1920 x 1080 Scanning and Analog and Parallel Digital Interfaces for Multiple Picture Rates, Society of Motion Picture and Television Engineers, White Plains, N.Y.

[F5] SMPTE 293M (2003): Standard for Television—720 x 483 Active Line at 59.94-Hz Progressive Scan Production, Digital Representation, Society of Motion Picture and Television Engineers, White Plains, N.Y.

[F6] SMPTE 296M (2001): Standard for Television—1280 x 720 Progressive Image Sample Structure, Analog and Digital Representation and Analog Interface, Society of Motion Picture and Television Engineers, White Plains, N.Y.

[F7] SMPTE 421M (2005): "VC-1 Compressed Video Bitstream Format and Decoding Process", Draft Standard, August 2005.

[F8] SMPTE RP227 (2005): "VC-1 Bitstream Transport Encodings", Draft Recommended Practice, August 2005.

[F9] CEA-708-B: "Digital Television (DTV) Closed Captioning," Consumer Electronics Association.

[F10] ITU-R BT.601-5 (1995): Encoding Parameters of Digital Television for Studios.

2.2 Informative References

[F11] ETSI TS 101 154 V1.7.1, Digital Video Broadcasting (DVB): Implementation Guidelines for the use of MPEG-2 Systems, Video and Audio in Satellite, Cable and Terrestrial Broadcasting Applications, Annex B, June 2005.

[F12] Digital Receiver Implementation Guidelines and Recommended Receiver Reaction to Aspect Ratio Signaling in Digital Video Broadcasting, Issue 1.2, August 2000, Digital TV Group.

[F13] ANSI/SCTE 21 2001 (formerly DVS 053): Standard for Carriage of NTSC VBI Data in Cable Digital Transport Streams, Society of Cable Telecommunications Engineers.

[F14] CEA-CEB-10-A: "EIA-708-B Implementation Guidance," December 2002, Consumer Electronics Association.

[F15] ITU-R BT. 709-5 (2002): Parameter values for the HDTV Standards for Production and International Programme Exchange

[F16] Digital Receiver Implementation Guidelines and Recommended Receiver Reaction to Aspect Ratio Signaling in Digital Video Broadcasting, Issue 1.2.1, February 2001, Digital TV Group.

3. COMPLIANCE NOTATION

As used in this document, "shall" denotes a mandatory provision of the standard. "Should" denotes a provision that is recommended but not mandatory. "May" denotes a feature whose presence does not preclude compliance, that may or may not be present at the option of the implementor.

4. POSSIBLE VIDEO INPUTS

While not required by this standard, there are certain television production standards, shown in Table F1, that define video formats that relate to compression formats specified by this standard.

Table F1 Standardized Video Input Formats

Video Standard	Active Lines	Active Samples/ Line
SMPTE 274M	1080	1920
SMPTE 296M	720	1280
SMPTE 293M	483	720
SMPTE 125M	483	720

The compression formats may be derived from one or more appropriate video input formats. It may be anticipated that additional video production standards will be developed in the future that extend the number of possible input formats.

5. SOURCE CODING SPECIFICATION

The VC-1 video compression algorithm shall conform to the Advanced Profile syntax specified in [F7]. The allowable parameters shall be bounded by the upper limits set forth in Level 4.

5.1 Carriage of VC-1 Elementary Streams in MPEG-2 Transport Streams

VC-1 elementary streams shall be carried in MPEG-2 Transport Streams as specified in [F8].

5.2 Constraints with Respect to VC-1 Advanced Profile

The following sub-sections list the allowed values for each of the SMPTE VC-1 syntactic elements which are restricted beyond the limits imposed by AP@L4.

In these tables conventional numbers denote decimal values, numbers preceded by 0x are to be interpreted as hexadecimal values and numbers within single quotes (e.g., '10010100') are to be interpreted as a string of binary digits.

5.2.1 Access Points

An Access Point is defined as an Access Unit in an Advanced Profile VC-1 bit stream at which a decoder can begin decoding video successfully. The Access Unit associated with an Access Point must contain a Sequence Header followed by an Entry Point Header structure.

5.2.2 Sequence Header Parameter Set Constraints

The constraints for the fields in a Sequence Header are listed in Table F2.

Table F2 Sequence Header Parameter Set Constraints

Sequence Parameter Set Syntactic Element	Allowed Value
PROFILE	See Table F3
LEVEL	See Table F3
CHROMAFORMAT	'1'
MAX_CODED_WIDTH	See Table F3
MAX_CODED_HEIGHT	See Table F3
TFCNTRFLAG	'1'
DISPLAY_EXT	'1'
ASPECT_RATIO_FLAG	'1'
ASPECT_RATIO	See Table F3
FRAMERATE_FLAG	'1'
FRAMERATEIND	'0'
FRAMERATENR	See Table F4
FRAMERATEDR	See Table F4
COLOR_FORMAT_FLAG	'1'
HRD_PARAM_FLAG	See Note 1 below

Note 1: The maximum size of the video elementary stream buffer shall comply with the value Rmax listed in Table 253 of [F7]. Because E-VSB transmission allows for multiple transmission rates, the sequence header may include zero, one or more Hypothetical Reference Decoder buffer model parameters in the form of a HRD_PARAM() structure in the sequence header and a HRD_FULLNESS() field in the Entry Point header. The HRD_PARAM_FLAG and HRD_NUM_LEAKY_BUCKETS shall be set accordingly, as specified in [F7].

5.2.3 Compression Format Constraints

Table F3 lists the allowed compression formats.

Table F3 Compression Format Constraints

Vertical Size	Horizontal Size	PROFILE	LEVEL	MAX_ CODED_ WIDTH	MAX_ CODED_ HEIGHT	ASPECT_ RATIO	Display Aspect Ratio	Allowed Frame Rates	Progressive interlaced
1080	1920	3	3	959	539	1	16:9	1,2,3,4	P
1080	1920	3	4 [1]	959	539	1	16:9	5,6	P
1080	1920	3	3	959	539	1	16:9	3,4	i
1080	1440	3	3	959	719	15 [2]	16:9	1,2,3,4	P
1080	1440	3	3	959	719	15 [2]	16:9	3,4	i
720	1280	3	2	359	639	1	16:9	1,2,3,4	P
720	1280	3	3	359	639	1	16:9	5,6	P
480	720	3	1	239	359	3	4:3	1,2,3,4	P
480	720	3	2	239	359	3	4:3	5,6	P
480	720	3	1	239	359	5	16:9	1,2,3,4	P
480	720	3	2	239	359	5	16:9	5,6	P
480	720	3	1	239	359	3	4:3	3,4	i
480	720	3	1	239	359	5	16:9	3,4	i
480	704	3	1	239	351	3	4:3	1,2,3,4	P
480	704	3	2	239	351	3	4:3	5,6	P
480	704	3	1	239	351	5	16:9	1,2,3,4	P
480	704	3	2	239	351	5	16:9	5,6	P
480	704	3	1	239	351	3	4:3	3,4	i
480	704	3	1	239	351	5	16:9	3,4	i
480	640	3	1	239	319	1	4:3	1,2,3,4	P
480	640	3	2	239	319	1	4:3	5,6	P
480	640	3	1	239	319	1	4:3	3,4	i
480	544	3	1	239	271	5	4:3	1	P
480	544	3	1	239	271	5	4:3	3	i
480	528	3	1	239	263	5	4:3	1	P
480	528	3	1	239	263	5	4:3	3	i
480	352	3	1	239	175	7	4:3	1	P
480	352	3	1	239	175	7	4:3	3	i
240	352	3	0	119	175	3	4:3	1	P
120	176	3	0	59	87	3	4:3	1	P

Legend:
frame rate: 1 = 23.976 Hz, 2 = 24 Hz, 3 = 29.97 Hz, 4 = 30 Hz, 5 = 59.94 Hz, 6 = 60 Hz
aspect_ratio_idc: 1 = 1:1 [square samples], 3 = 10:11, 5 = 40:33, 7 = 20:11, 11 = 15:11

Note:
For vertical sizes of 1080 and 120, 1088, and 128 lines (respectively) are actually coded in order to satisfy the VC-1 requirement that the coded vertical size be a multiple of 16 (progressive scan) or 32 (interlaced scan).

Footnotes:
[1] Use of this format may be constrained by service_type (see A/65).
[2] In addition, aspect_width shall be equal to 4 and aspect_height shall be equal to 3.

For each frame rate shown in Table F3, the values for FRAMERATENR and FRAMERATEDR shall be as shown in Table F4.

Table F4 Frame rate Parameter Constraints

frame rate	FRAMERATENR	FRAMERATEDR
23.98 Hz	1	2
24 Hz	1	1
29.97 Hz	3	2
30 Hz	3	1
59.94 Hz	5	2
60 Hz	5	1
Legend: FRAMERATENR: 1 = 24*1000, 3 = 30*1000, 5 = 60*1000 FRAMERATEDR: 1 = 1000, 2 = 1001		

5.3 Low Delay and Still Picture Modes

5.3.1 Low Delay Mode

Low delay mode shall follow the variable-delay mode of the VC-1 Hypothetical Reference Decoder (HRD) model as defined in Annex C of the VC-1 Specification [F7]. This mode of operation is signaled when there is no HRD parameter signaled in the sequence header; i.e., when HRD_PARAM_FLAG is equal to '0'.

5.3.2 Still Picture Mode

Still Picture Mode is supported and is governed by the conventional MPEG-2 Systems principles designed to support this mechanism [F1]. In particular, still picture mode is characterized by:

- A VC-1 sequence header followed by an optional Entry Point Header followed by an I-frame followed by an End-of-Sequence code.
- The presentation of the next frame is greater or equal to the value of the Presentation Time Stamp for the current frame plus 2 frame periods. In particular, the constraint that consecutive Presentation Time Stamps be less than 700 milliseconds apart is not applicable for VC-1 still pictures.
- The time elapsed between the instant when the first byte of the I-frame enters the video elementary stream and its Decoding Time Stamp is smaller or equal to 60 seconds.

5.4 Bit Stream Specifications for User Data and Display Metadata

This section covers the user data and display metadata parts of the video syntax. These data are inserted at the sequence, entry point or frame level. The syntax used for the insertion of closed captioning in frame user data is described. Much of the following material is duplicated or modified from material contained in Annex A, Section 5.2, and is provided here for the convenience of the reader.

5.4.1 User Data Syntax

Table F5 describes the user data syntax. The table is consistent with the rule established in Annex F of [F7].

Table F5 User Data Syntax[1]

Syntax	No. of Bits	Format
VC1_user_data( ) {		
VC1_user_data_start_code	32	bslbf
ATSC_identifier	32	bslbf
user_data_type_code	8	uimsbf
if (user_data_type_code == '0x03')		
cc_data()[2]		
else {		
while (nextbits() != '0000 0000 0000 0000 0000 0001') {		
ATSC_reserved_user_data	8	
}		
flushing_byte	8	0x80
next_start_code()		
}		

In accordance with the bit stream syntax in Table F5, more than one user data construct may follow any given header. However, no more than one user data construct using the same user_data_type_code value shall be in any given header.

Note that user data with a 32-bit field following VC1_user_data_start_code having a value other than ATSC_identifier may be present in an ATSC-compliant video bit stream. As an example, the afd_identifier (value 0x44544731) is defined for use in ATSC video Elementary Streams (see Section 5.3.3). Receiving devices are expected to process this field and use it to determine the syntax and semantics of the user data construct to follow.

Receiving devices are expected to silently discard any unrecognized video user data encountered in the video bit stream. For example, if an unrecognized 32-bit identifier is seen following the VC1_user_data_start_code, or an unrecognized 8-bit user_data_type_code is seen following the ATSC_identifier, data should be discarded until another start code is seen.

5.4.2 ATSC Picture User Data Semantics

VC1_user_data_start_code – This is set to 0x0000 011D.

ATSC_identifier – This is a 32 bit code that indicates that the video user data conforms to this specification. The value ATSC_identifier shall be 0x4741 3934.

user_data_type_code – An 8-bit value that identifies the type of ATSC user data to follow. Value 0x03 indicates cc_data(), and other values are either in use in other standards or are reserved for future use.

cc_data() – A data structure defined in Table F6.

ATSC_reserved_user_data – Reserved for use by ATSC or used by other standards.

[1] Shaded cells in this table indicate syntactic and semantic additions to [F7].

[2] cc_data() shall only be included in a frame-level VC1_user_data() structure (see Section 5.4.2.1).

5.4.2.1 Captioning Data

Captioning data shall only be included in a frame-level VC1_user_data() structure. The syntax of captioning data is described in Annex A of this document. It is reproduced in Table F6 below for the convenience of the reader.

Table F6 Captioning Data Syntax

Syntax	No. of Bits	Format
cc_data() {		
reserved	1	'1'
process_cc_data_flag	1	bslbf
additional_data_flag	1	bslbf
cc_count	5	uimsbf
reserved	8	'1111 1111'
for (i=0 ; i < cc_count ; i++) {		
marker_bits	5	'1111 1'
cc_valid	1	bslbf
cc_type	2	bslbf
cc_data_1	8	bslbf
cc_data_2	8	bslbf
}		
marker_bits	8	'1111 1111'
if (additional_data_flag) {		
while (nextbits() != '0000 0000 0000 0000 0000 0001') {		
additional_cc_data		
}		
}		
}		

process_cc_data_flag – This flag is set to indicate whether it is necessary to process the cc_data. If it is set to 1, the cc_data has to be parsed and its meaning has to be processed. When it is set to 0, the cc_data can be discarded.

additional_data_flag – This flag is set to 1 to indicate the presence of additional user data.

cc_count: – This 5-bit integer indicates the number of closed caption constructs following this field. It can have values 0 through 31. The value of cc_count shall be set according to the frame rate and coded picture structure (field or frame) such that a fixed bandwidth of 9600 bits per second is maintained for the closed caption payload data. Sixteen (16) bits of closed caption payload data are carried in each pair of the fields cc_data_1 and cc_data_2.

cc_valid – This flag is set to '1' to indicate that the two closed caption data bytes that follow are valid. If set to '0' the two data bytes are invalid, as defined in [F9].

cc_type – Denotes the type of the two closed caption data bytes that follow, as defined in [F9].

cc_data_1 – The first byte of a closed caption data pair as defined in [F9].

cc_data_2 – The second byte of a closed caption data pair as defined in [F9].

additional_cc_data – Reserved for future ATSC definition.

5.4.3 Bar Data

Bar data shall be conveyed in the VC-1 elementary stream by means of the pan scan information in the picture header. The PANSCAN_FLAG value '1' in the entry point header indicates bar data is present in the VC-1 bitstream. The set of pan scan values, PS_HOFFSET, PS_VOFFSET, PS_WIDTH, and PS_HEIGHT, remains in effect until the PS_PRESENT flag in a subsequent picture header is '1' or until the PANSCAN_FLAG value in a subsequent entry point header is reset to '0', whichever comes first.

The reference point for the pan scan fields, PS_HOFFSET, PS_VOFFSET, PS_WIDTH, and PS_HEIGHT, is the upper left corner of the video frame. Designation of line numbers for calculating the coordinates of the window within the frame shall conform to the applicable standards as listed in Table F7.

Table F7 Line Number Designation

Video Format	Applicable Standard
480 Interlaced (analog)	SMPTE 170M [F3]
480 Interlaced 4:3 (digital)	SMPTE 125M [F2]
480 Progressive	SMPTE 293M [F5]
720 Progressive	SMPTE 296M [F6]
1080 Interlaced	SMPTE 274M [F4]
1080 Progressive	SMPTE 274M [F4]

5.4.4 Active Format Description Data

Active Format Description (AFD) is described in Annex A of this document. The text from Annex A is reproduced in this section for the convenience of the reader. AFD should be included in video user data whenever the rectangular picture area containing useful information does not extend to the full height or width of the coded frame. AFD data may also be included in user data when the rectangular picture area containing useful information extends to the full height and width of the coded frame.

When present, the AFD shall be carried, using the syntax and semantics defined in [F11], in the user data of the video elementary stream. After each sequence start (and repeat sequence start) the default aspect ratio of the area of interest shall be that signalled by the sequence header and sequence display extension parameters. After introduction, an AFD shall remain in effect until the next sequence start or until another AFD is introduced. Receivers should interpret the absence of AFD in a sequence start to mean the active format is the same as the coded frame, corresponding to AFD value '1000' (see Table F9).

> Note: The AFD syntax as shown here is syntactically identical to that specified in ETSI TS 101 154 V1.7.1 [F11], and is reprinted here with permission.

5.4.4.1 AFD Syntax

Table F8 describes the syntax of the Active Format Description.

Table F8 Active Format Description Syntax

Syntax	No. of Bits	Format
VC1_user_data_start_code	32	bslbf
afd_identifier	32	bslbf
zero	1	'0'
active_format_flag	1	bslbf
flag_bits	6	'00 0001'
if (active_format_flag == 1) {		
reserved	4	'1111'
active_format	4	bslbf
}		

5.4.4.2 AFD Semantics

VC1_user_data_start_code – This 32-bit field shall be set to 0x0000011D to indicate the beginning of a user data structure in the VC-1 elementary stream.

afd_identifier – A 32-bit field that identifies that the syntax of the user data is Active Format Description. Its value is 0x44544731.

active_format_flag – A 1 bit flag. A value of '1' indicates that an active format is described in this data structure.

flag_bits – A 6 bit field with a defined value of '00 0001'.

active_format – A 4 bit field describing the "area of interest" in terms of its aspect ratio within the coded frame as defined in [F7].

> The active_format is used by the decoder in conjunction with the sample aspect ratio signaled in a VC-1 elementary stream by means of the ASPECT_RATIO field in the sequence header.

The combination of sample aspect ratio and active_format allows the decoder to identify whether the "area of interest" is the whole of the frame (e.g., source aspect ratio 16:9, active_format 16:9 center), a letterbox within the frame (e.g., source aspect ratio 4:3, active_format 16:9 center), or a "pillarbox" within the frame (e.g., source aspect ratio 16:9, active_format 4:3 center).

Table F9 defines the coding of the active_format field that shall be used.

Table F9 Active Format

active_format	Description	
	4:3 coded frames	**16:9 coded frames**
'0000' – '0001'	Reserved	Reserved
'0010' – '0100"	Not recommended	Not recommended
'0101' – '0111'	Reserved	Reserved
'1000'	4:3 full frame image	16:9 full frame image
'1001'	4:3 full frame image	4:3 pillarbox image
'1010'	16:9 letterbox image	16:9 full frame image
'1011'	14:9 letterbox image	14:9 pillarbox image
'1100'	Reserved	Reserved
'1101'	4:3 full frame image, alternative 14:9 center	4:3 pillarbox image, alternative 14:9 center
'1110'	16:9 letterbox image, alternative 14:9 center	16:9 full frame image, alternative 14:9 center
'1111'	16:9 letterbox image, alternative 4:3 center	16:9 full frame image, alternative 4:3 center

Use of '0010', '0011' or '0100' is not recommended. Illustrations of the various values of active_format may be found in ETSI-TS 101 154 Table B.3 [F11]. Users are strongly encouraged to consult this reference.

5.4.4.3 Recommended Receiver Response to AFD

Receiving device designers are strongly encouraged to study the suite of Consumer Electronics Association (CEA) bulletins which contain the most recent recommendations regarding the processing of AFD. In several instances, a variety of design choices are possible when processing a given AFD value for display and the recommendation identifies one preferred method.

5.4.5 Relationship Between Bar Data and AFD (Informative)

Any combination of Active Format Description and bar data may be present in video user data (either, neither, or both). Note that AFD data may not always exactly match bar data because AFD only deals with 4:3, 14:9, and 16:9 aspect ratios while bar data can accurately represent nearly any aspect ratio. Whenever bar data is present, it should be assumed to be exact and a refinement of the AFD signaling.

-- End of document --