

Doc. A/53
16 Sep 1995

Doc. A/53A
6 April 2001

Doc. A/53B
7 August 2001

ATSC Standard: Digital Television Standard, Revision B, with Amendment 1

Advanced Television Systems Committee

1750 K Street, N.W.

Suite 1200

Washington, D.C. 20006

www.atsc.org

The Advanced Television Systems Committee (ATSC), is an international, non-profit membership organization developing voluntary standards for the entire spectrum of advanced television systems.

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 160 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.

Table of Contents

1. SCOPE AND DOCUMENTATION STRUCTURE	9
1.1 Documentation Structure	9
2. REFERENCES.....	9
3. DEFINITIONS	9
3.1 Compliance Notation	10
3.2 Treatment of Syntactic Elements	10
3.3 Terms Employed	10
3.4 Symbols, Abbreviations, and Mathematical Operators	16
3.4.1 Arithmetic Operators	16
3.4.2 Logical Operators	17
3.4.3 Relational Operators	17
3.4.4 Bitwise Operators	17
3.4.5 Assignment	17
3.4.6 Mnemonics	17
3.4.7 Constants	17
3.4.8 Method of Describing Bit Stream Syntax	18
3.4.8.1 Definition of bytealigned function	19
3.4.8.2 Definition of nextbits function	19
3.4.8.3 Definition of next_start_code function	19
4. BACKGROUND.....	20
4.1 Advisory Committee on Advanced Television Service (ACATS)	20
4.2 Digital HDTV Grand Alliance (Grand Alliance)	21
4.3 Organization for Documenting the Digital Television Standard	21
4.4 Principles for Documenting the Digital Television Standard	22
5. SYSTEM OVERVIEW	23
5.1 System Block Diagram	23
 Annex A: Video System Characteristics (Normative)	
1. SCOPE.....	27
2. REFERENCES.....	27
2.1 Normative References	27
2.2 Informative References	27

3. COMPLIANCE NOTATION	27
4. POSSIBLE VIDEO INPUTS.....	27
5. SOURCE CODING SPECIFICATION.....	28
5.1 Constraints with Respect to ISO/IEC 13818-2 Main Profile	28
5.1.1 Sequence Header Constraints	28
5.1.2 Compression Format Constraints	29
5.1.3 Sequence Extension Constraints	29
5.1.4 Sequence Display Extension Constraints	30
5.1.5 Picture Header Constraints	30
5.1.6 Picture Coding Constraints	30
5.2 Bit Stream Specifications Beyond MPEG-2	30
5.2.1 Picture Extension and User Data Syntax	30
5.2.2 Picture User Data Syntax	31
5.2.3 Picture User Data Semantics	32

Annex B: Audio System Characteristics (Normative)

1. SCOPE.....	34
2. NORMATIVE REFERENCES	34
3. COMPLIANCE NOTATION	34
4. SYSTEM OVERVIEW	34
5. SPECIFICATION	35
5.1 Constraints With Respect to ATSC Standard A/52A	35
5.2 Sampling Frequency	36
5.3 Bit Rate	36
5.4 Audio Coding Modes	36
5.5 Dialogue Level	36
5.6 Dynamic Range Compression	36
5.7 STD Audio Buffer Size	36
6. MAIN AND ASSOCIATED SERVICES.....	36
6.1 Summary of Service Types	37
6.2 Complete Main Audio Service (CM)	37
6.3 Main Audio Service, Music and Effects (ME)	37
6.4 Visually Impaired (VI)	38
6.5 Hearing Impaired (HI)	38
6.6 Dialogue (D)	38
6.7 Commentary (C)	39
6.8 Emergency (E)	39

6.9	Voice-Over (V0)	39
7.	AUDIO ENCODER INTERFACES	40
7.1	Audio Encoder Input Characteristics	40
7.2	Audio Encoder Output Characteristics	40
Annex C: Service Multiplex & Transport System Characteristics (Normative)		
1.	SCOPE	41
2.	NORMATIVE REFERENCES	41
3.	COMPLIANCE NOTATION	41
4.	SYSTEM OVERVIEW	41
5.	SPECIFICATION	43
5.1	MPEG-2 Systems Standard	43
5.1.1	Video T-STD	43
5.1.2	Audio T-STD	43
5.2	Registration Descriptor	43
5.2.1	Program Identifier	43
5.2.2	Audio Elementary Stream Identifier	43
5.3	Audio Constraints	43
5.4	Constraints on PSI	44
5.5	PES Constraints	44
5.5.1	Video PES Constraints	44
5.5.2	Audio PES Constraints	45
5.6	Services and Features	45
5.6.1	System Information and Program Guide	45
5.6.1.1	System information and program guide PID	45
5.6.1.2	System information and program guide STD model	46
5.6.2	Specification of Private Data Services	46
5.6.2.1	Verification model	46
5.6.2.1.1	Syntax and semantics	46
5.6.2.1.2	Ancillary service target decoder (ASTD)	46
5.6.2.2	Stream type and PMT descriptors	47
5.6.2.2.1	Stream type	47
5.6.2.2.2	PMT descriptors	47
5.7	Assignment of Identifiers	47
5.7.1	Stream Type	47
5.7.2	Descriptors	48
5.7.2.1	AC-3 audio descriptor	48
5.7.2.2	Program smoothing buffer descriptor	48
5.8	Extensions to the MPEG-2 Systems Specification	48

5.8.1	Scrambling Control	48
6.	FEATURES OF 13818-1 NOT SUPPORTED BY THIS STANDARD	49
6.1	Program Streams	49
6.2	Still Pictures	49
7.	TRANSPORT ENCODER INTERFACES AND BIT RATES	49
7.1	Transport Encoder Input Characteristics	49
7.2	Transport Output Characteristics	49

Annex D: RF/Transmission System Characteristics (Normative)

1.	SCOPE	50
2.	NORMATIVE REFERENCESE	50
3.	COMPLIANCE NOTATION	50
4.	TRANSMISSION CHARACTERISTICS FOR TERRESTRIAL BROADCAST	50
4.1	Channel Error Protection and Synchronization	53
4.1.1	Prioritization	53
4.1.2	Data Randomizer	53
4.1.3	Reed-Solomon Encoder	54
4.1.4	Interleaving	55
4.1.5	Trellis Coding	56
4.1.6	Data Segment Sync	60
4.1.7	Data Field Sync	61
4.1.7.1	Sync	61
4.1.7.2	PN511	61
4.1.7.3	PN63	62
4.1.7.4	VSB mode	62
4.1.7.5	Reserved	63
4.1.7.6	Precode	63
4.2	Modulation	63
4.2.1	Bit-to-Symbol Mapping	63
4.2.2	Pilot Addition	63
4.2.3	8 VSB Modulation Method	64
5.	TRANSMISSION CHARACTERISTICS FOR HIGH DATA RATE MODE	64
5.1	Channel Error Protection and Synchronization	66
5.1.1	Prioritization	66
5.1.2	Data Randomizer	66
5.1.3	Reed-Solomon Encoder	66
5.1.4	Interleaving	66
5.1.5	Data Segment Sync	66
5.1.6	Data Field Sync	66

5.2 Modulation	67
5.2.1 Bit-to-Symbol Mapping	67
5.2.2 Pilot Addition	67
5.2.3 16 VSB Modulation Method	67

Annex E: Receiver Characteristics (Informative)

1. SCOPE	68
2. REFERENCES TO EXISTING OR EMERGING STANDARDS	68
3. COMPLIANCE NOTATION	68
4. STATUS OF RECEIVER STANDARDIZATION ACTIVITIES	68
4.1 Tuner Performance	68
4.1.1 Noise Figure	69
4.1.2 Channelization Plan for Broadcast and Cable	69
4.1.3 Direct Pickup	69
4.2 Transport	69
4.3 Decoder Interface	69
4.4 Digital Data Interface	69
4.5 Conditional Access Interface	69
4.6 Closed Captioning	70
5. RECEIVER FUNCTIONALITY	70
5.1 Video	70
5.2 Audio	70
Amendment 1 to A/53B	71

Index of Tables and Figures

Table 3.1 Next Start Code	19
Table A1 Standardized Video Input Formats	28
Table A2 Sequence Header Constraints	28
Table A3 Compression Format Constraints	29
Table A4 Sequence Extension Constraints	29
Table A5 Sequence Display Extension Constraints	30
Table A6 Picture Extension and User Data Syntax	31
Table A7 Picture User Data Syntax	32
Table B1 Audio Constraints	35
Table B2 Audio Service Types	37
Table C1 Transport Scrambling Control Field	48
Table D1 Interleaving Sequence	58
Table D2 Byte to Symbol Conversion, Multiplexing of Trellis Encoders	59
Figure 5.1 ITU-R digital terrestrial television broadcasting model.	24
Figure 5.2 High level view of encoding equipment.	25
Figure B1 Audio subsystem in the digital television system.	35
Figure C1 Sample organization of functionality in a transmitter-receiver pair for a single program.	42
Figure C2 Ancillary service target decoder.	47
Figure D1 VSB transmitter.	51
Figure D2 VSB data frame.	52
Figure D3 VSB channel occupancy (nominal).	53
Figure D4 Randomizer polynomial.	54
Figure D5 Reed-Solomon (207,187) $t=10$ parity generator polynomial.	55
Figure D6 Convolutional interleaver (byte shift register illustration).	56
Figure D7 8 VSB trellis encoder, precoder, and symbol mapper.	57
Figure D8 Trellis code interleaver.	57
Figure D9 8 VSB data segment.	60
Figure D10 VSB data field sync.	61
Figure D11 Field sync PN sequence generators.	62
Figure D12 Nominal VSB system channel response (linear phase raised cosine Nyquist filter).	64
Figure D13 16 VSB data segment.	65
Figure D14 16 VSB transmitter.	65
Figure D15 16 VSB mapper.	66

ATSC Digital Television Standard

1. SCOPE AND DOCUMENTATION STRUCTURE

The Digital Television Standard describes the system characteristics of the U. S. advanced television (ATV) system. The document and its normative annexes provide detailed specification of the parameters of the system including the video encoder input scanning formats and the pre-processing and compression parameters of the video encoder, the audio encoder input signal format and the pre-processing and compression parameters of the audio encoder, the service multiplex and transport layer characteristics and normative specifications, and the VSB RF/Transmission subsystem.

1.1 Documentation Structure

The documentation of the Digital Television Standard consists of this document which provides a general system overview, a list of reference documents, and sections relating to the system as a whole. The system is modular in concept and the specifications for each of the modules are provided in the appropriate annex.

2. REFERENCES

Normative references may be found in each normative Annex. The Digital Television Standard is based on the ISO/IEC MPEG-2 Video Standard, the Digital Audio Compression (AC-3) Standard, and the ISO/IEC MPEG-2 Systems Standard. Those references are listed here for the convenience of the reader. In addition, a guide to the use of the Digital Television Standard is listed.

ATSC Standard A/52 (1995), Digital Audio Compression (AC-3).

ATSC Document A/54 (1995), Guide to the Use of the ATSC Digital Television Standard.

ISO/IEC IS 13818-1, International Standard (1996), MPEG-2 Systems.

ISO/IEC 13818-1: 1996/Cor. 1: 1997 (E) Technical Corrigendum 1.

ISO/IEC 13818-1: 1996/Amd. 1: 1997 (E) Amendment 1.

ISO/IEC 13818-1: 1996/Amd. 2: 1997 (E) Amendment 2.

ISO/IEC 13818-1: 1996/Amd. 3: 1997 (E) Amendment 3.

ISO/IEC 13818-1: 1996/Amd. 4: 1997 (E) Amendment 4.

ISO/IEC IS 13818-2, International Standard (1996), MPEG-2 Video.

ISO/IEC 13818-2: 1996/Cor. 1: 1997 (E) MPEG-2 Video Technical Corrigendum 1.

ISO/IEC 13818-2: 1996/Cor. 2: 1997 (E) MPEG-2 Video Technical Corrigendum 2.

3. DEFINITIONS

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 shall be used. Where an abbreviation is not covered by IEEE practice, or industry practice differs from IEEE practice, then the abbreviation in question will be described in Section 3.4 of this

document. Many of the definitions included therein are derived from definitions adopted by MPEG.

3.1 Compliance Notation

As used in this document, “shall” or “will” 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.

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*).

3.3 Terms Employed

For the purposes of the Digital Television Standard, the following definition of terms apply:

ACATS: Advisory Committee on Advanced Television Service.

access unit: A coded representation of a presentation unit. In the case of audio, an access unit is the coded representation of an audio frame. In the case of video, an access unit includes all the coded data for a picture, and any stuffing that follows it, up to but not including the start of the next access unit. If a picture is not preceded by a *group_start_code* or a *sequence_header_code*, the access unit begins with a the picture start code. If a picture is preceded by a *group_start_code* and/or a *sequence_header_code*, the access unit begins with the first byte of the first of these start codes. If it is the last picture preceding a *sequence_end_code* in the bit stream all bytes between the last byte of the coded picture and the *sequence_end_code* (including the *sequence_end_code*) belong to the access unit.

A/D: Analog to digital converter.

AES: Audio Engineering Society.

anchor frame: A video frame that is used for prediction. I-frames and P-frames are generally used as anchor frames, but B-frames are never anchor frames.

ANSI: American National Standards Institute.

Asynchronous Transfer Mode (ATM): A digital signal protocol for efficient transport of both constant-rate and bursty information in broadband digital networks. The ATM digital stream consists of fixed-length packets called “cells,” each containing 53 8-bit bytes—a 5-byte header and a 48-byte information payload.

ATEL: Advanced Television Evaluation Laboratory.

ATM: See *asynchronous transfer mode*.

ATTC: Advanced Television Test Center.

ATV: The U. S. advanced television system.

bidirectional pictures or **B-pictures** or **B-frames:** Pictures that use both future and past pictures as a reference. This technique is termed *bidirectional prediction*. B-pictures provide the

most compression. B-pictures do not propagate coding errors as they are never used as a reference.

bit rate: The rate at which the compressed bit stream is delivered from the channel to the input of a decoder.

block: A block is an 8-by-8 array of pel values or DCT coefficients representing luminance or chrominance information.

bps: Bits per second.

byte-aligned: A bit in a coded bit stream is byte-aligned if its position is a multiple of 8-bits from the first bit in the stream.

CDTV: See *conventional definition television*.

channel: A digital medium that stores or transports a digital television stream.

coded representation: A data element as represented in its encoded form.

compression: Reduction in the number of bits used to represent an item of data.

constant bit rate: Operation where the bit rate is constant from start to finish of the compressed bit stream.

conventional definition television (CDTV): This term is used to signify the *analog* NTSC television system as defined in ITU-R Recommendation 470. See also *standard definition television* and ITU-R Recommendation 1125.

CRC: The cyclic redundancy check to verify the correctness of the data.

D-frame: Frame coded according to an MPEG-1 mode which uses DC coefficients only.

data element: An item of data as represented before encoding and after decoding.

DCT: See *discrete cosine transform*.

decoded stream: The decoded reconstruction of a compressed bit stream.

decoder: An embodiment of a decoding process.

decoding (process): The process defined in the Digital Television Standard that reads an input coded bit stream and outputs decoded pictures or audio samples.

decoding time-stamp (DTS): A field that may be present in a PES packet header that indicates the time that an access unit is decoded in the system target decoder.

digital storage media (DSM): A digital storage or transmission device or system.

discrete cosine transform: A mathematical transform that can be perfectly undone and which is useful in image compression.

DSM-CC: Digital storage media command and control.

DSM: Digital storage media.

DTS: See *decoding time-stamp*.

DVCR: Digital video cassette recorder

ECM: See *entitlement control message*.

editing: A process by which one or more compressed bit streams are manipulated to produce a new compressed bit stream. Conforming edited bit streams are understood to meet the requirements defined in the Digital Television Standard.

elementary stream (ES): A generic term for one of the coded video, coded audio, or other coded bit streams. One elementary stream is carried in a sequence of PES packets with one and only one `stream_id`.

elementary stream clock reference (ESCR): A time stamp in the PES stream from which decoders of PES streams may derive timing.

EMM: See *entitlement management message*.

encoder: An embodiment of an encoding process.

encoding (process): A process that reads a stream of input pictures or audio samples and produces a valid coded bit stream as defined in the Digital Television Standard.

entitlement control message (ECM): Entitlement control messages are private conditional access information which specify control words and possibly other stream-specific, scrambling, and/or control parameters.

entitlement management message (EMM): Entitlement management messages are private conditional access information which specify the authorization level or the services of specific decoders. They may be addressed to single decoders or groups of decoders.

entropy coding: Variable length lossless coding of the digital representation of a signal to reduce redundancy.

entry point: Refers to a point in a coded bit stream after which a decoder can become properly initialized and commence syntactically correct decoding. The first transmitted picture after an entry point is either an I-picture or a P-picture. If the first transmitted picture is not an I-picture, the decoder may produce one or more pictures during acquisition.

ES: See *elementary stream*.

ESCR: See *elementary stream clock reference*.

event: An event is defined as a collection of elementary streams with a common time base, an associated start time, and an associated end time.

field: For an interlaced video signal, a “field” is the assembly of alternate lines of a frame. Therefore, an interlaced frame is composed of two fields, a top field and a bottom field.

forbidden: This term, when used in clauses defining the coded bit stream, indicates that the value shall never be used. This is usually to avoid emulation of start codes.

FPLL: Frequency and phase locked loop.

frame: A frame contains lines of spatial information of a video signal. For progressive video, these lines contain samples starting from one time instant and continuing through successive lines to the bottom of the frame. For interlaced video a frame consists of two fields, a top field and a bottom field. One of these fields will commence one field later than the other.

GOP: See *group of pictures*.

Group of pictures (GOP): A group of pictures consists of one or more pictures in sequence.

HDTV: See *high definition television*.

high definition television (HDTV): High definition television has a resolution of approximately twice that of conventional television in both the horizontal (H) and vertical (V) dimensions and a picture aspect ratio (H × V) of 16:9. ITU-R Recommendation 1125 further defines “HDTV quality” as the delivery of a television picture which is subjectively identical with the interlaced HDTV studio standard.

high level: A range of allowed picture parameters defined by the MPEG-2 video coding specification which corresponds to high definition television.

Huffman coding: A type of source coding that uses codes of different lengths to represent symbols which have unequal likelihood of occurrence.

IEC: International Electrotechnical Commission.

intra-coded pictures or **I-pictures** or **I-frames:** Pictures that are coded using information present only in the picture itself and not depending on information from other pictures. I-pictures provide a mechanism for random access into the compressed video data. I-pictures employ transform coding of the pel blocks and provide only moderate compression.

ISO: International Organization for Standardization.

ITU: International Telecommunication Union.

JEC: Joint Engineering Committee of EIA and NCTA.

layer: One of the levels in the data hierarchy of the video and system specification.

level: A range of allowed picture parameters and combinations of picture parameters.

macroblock: In the advanced television system, a macroblock consists of four blocks of luminance and one each C_r and C_b block.

main level: A range of allowed picture parameters defined by the MPEG-2 video coding specification with maximum resolution equivalent to ITU-R Recommendation 601.

main profile: A subset of the syntax of the MPEG-2 video coding specification that is expected to be supported over a large range of applications.

Mbps: 1,000,000 bits per second.

motion vector: A pair of numbers which represent the vertical and horizontal displacement of a region of a reference picture for prediction.

MP@HL: Main profile at high level.

MP@ML: Main profile at main level.

MPEG: Refers to standards developed by the ISO/IEC JTC1/SC29 WG11, *Moving Picture Experts Group*. MPEG may also refer to the Group.

MPEG-1: Refers to ISO/IEC standards 11172-1 (Systems), 11172-2 (Video), 11172-3 (Audio), 11172-4 (Compliance Testing), and 11172-5 (Technical Report).

MPEG-2: Refers to ISO/IEC standards 13818-1 (Systems), 13818-2 (Video), 13818-3 (Audio), 13818-4 (Compliance).

pack: A pack consists of a pack header followed by zero or more packets. It is a layer in the system coding syntax.

packet data: Contiguous bytes of data from an elementary data stream present in the packet.

packet identifier (PID): A unique integer value used to associate elementary streams of a program in a single or multi-program transport stream.

packet: A packet consists of a header followed by a number of contiguous bytes from an elementary data stream. It is a layer in the system coding syntax.

padding: A method to adjust the average length of an audio frame in time to the duration of the corresponding PCM samples, by continuously adding a slot to the audio frame.

payload: Payload refers to the bytes which follow the header byte in a packet. For example, the payload of a transport stream packet includes the PES_packet_header and its PES_packet_data_bytes or pointer_field and PSI sections, or private data. A PES_packet_payload, however, consists only of PES_packet_data_bytes. The transport stream packet header and adaptation fields are not payload.

PCR: See *program clock reference*.

pel: See *pixel*.

PES packet header: The leading fields in a PES packet up to but not including the PES_packet_data_byte fields where the stream is not a padding stream. In the case of a padding stream, the PES packet header is defined as the leading fields in a PES packet up to but not including the padding_byte fields.

PES packet: The data structure used to carry elementary stream data. It consists of a packet header followed by PES packet payload.

PES Stream: A PES stream consists of PES packets, all of whose payloads consist of data from a single elementary stream, and all of which have the same stream_id.

PES: An abbreviation for packetized elementary stream.

picture: Source, coded, or reconstructed image data. A source or reconstructed picture consists of three rectangular matrices representing the luminance and two chrominance signals.

PID: See *packet identifier*.

pixel: “Picture element” or “pel.” A pixel is a digital sample of the color intensity values of a picture at a single point.

predicted pictures or P-pictures or P-frames: Pictures that are coded with respect to the nearest *previous* I or P-picture. This technique is termed *forward prediction*. P-pictures provide more compression than I-pictures and serve as a reference for future P-pictures or B-pictures. P-pictures can propagate coding errors when P-pictures (or B-pictures) are predicted from prior P-pictures where the prediction is flawed.

presentation time-stamp (PTS): A field that may be present in a PES packet header that indicates the time that a presentation unit is presented in the system target decoder.

presentation unit (PU): A decoded audio access unit or a decoded picture.

profile: A defined subset of the syntax specified in the MPEG-2 video coding specification

program clock reference (PCR): A time stamp in the transport stream from which decoder timing is derived.

program element: A generic term for one of the elementary streams or other data streams that may be included in the program.

program specific information (PSI): PSI consists of normative data which is necessary for the demultiplexing of transport streams and the successful regeneration of programs.

program: A program is a collection of program elements. Program elements may be elementary streams. Program elements need not have any defined time base; those that do have a common time base and are intended for synchronized presentation.

PSI: See *program specific information*.

PTS: See *presentation time-stamp*.

PU: See *presentation unit*.

quantizer: A processing step which intentionally reduces the precision of DCT coefficients.

random access: The process of beginning to read and decode the coded bit stream at an arbitrary point.

reserved: This term, when used in clauses defining the coded bit stream, indicates that the value may be used in the future for Digital Television Standard extensions. Unless otherwise specified within this Standard, all reserved bits shall be set to “1”.

SCR: See *system clock reference*.

scrambling: The alteration of the characteristics of a video, audio, or coded data stream in order to prevent unauthorized reception of the information in a clear form. This alteration is a specified process under the control of a conditional access system.

SDTV: See *standard definition television*.

slice: A series of consecutive macroblocks.

SMPTE: Society of Motion Picture and Television Engineers.

source stream: A single, non-multiplexed stream of samples before compression coding.

splicing: The concatenation performed on the system level or two different elementary streams. It is understood that the resulting stream must conform totally to the Digital Television Standard.

standard definition television (SDTV): This term is used to signify a *digital* television system in which the quality is approximately equivalent to that of NTSC. This equivalent quality may be achieved from pictures sourced at the 4:2:2 level of ITU-R Recommendation 601 and subjected to processing as part of the bit rate compression. The results should be such that when judged across a representative sample of program material, subjective equivalence with NTSC is achieved. Also called standard digital television. See also *conventional definition television* and ITU-R Recommendation 1125.

start codes: 32-bit codes embedded in the coded bit stream that are unique. They are used for several purposes including identifying some of the layers in the coding syntax. Start codes consist of a 24 bit prefix (0x000001) and an 8 bit `stream_id`.

STD input buffer: A first-in, first-out buffer at the input of a system target decoder for storage of compressed data from elementary streams before decoding.

STD: See *system target decoder*.

still picture: A coded still picture consists of a video sequence containing exactly one coded picture which is intra-coded. This picture has an associated PTS and the presentation time of succeeding pictures, if any, is later than that of the still picture by at least two picture periods.

system clock reference (SCR): A time stamp in the program stream from which decoder timing is derived.

system header: The system header is a data structure that carries information summarizing the system characteristics of the Digital Television Standard multiplexed bit stream.

system target decoder (STD): A hypothetical reference model of a decoding process used to describe the semantics of the Digital Television Standard multiplexed bit stream.

time-stamp: A term that indicates the time of a specific action such as the arrival of a byte or the presentation of a presentation unit.

TOV: Threshold of visibility.

Transport Stream packet header: The leading fields in a Transport Stream packet up to and including the `continuity_counter` field.

variable bit rate: Operation where the bit rate varies with time during the decoding of a compressed bit stream.

VBV: See *video buffering verifier*.

Video buffering verifier (VBV): A hypothetical decoder that is conceptually connected to the output of an encoder. Its purpose is to provide a constraint on the variability of the data rate that an encoder can produce.

video sequence: A video sequence is represented by a sequence header, one or more groups of pictures, and an `end_of_sequence` code in the data stream.

8 VSB: Vestigial sideband modulation with 8 discrete amplitude levels.

16 VSB: Vestigial sideband modulation with 16 discrete amplitude levels.

3.4 Symbols, Abbreviations, and Mathematical Operators

The symbols, abbreviations, and mathematical operators used to describe the Digital Television Standard are those adopted for use in describing MPEG-2 and are similar to those used in the “C” programming language. However, integer division with truncation and rounding are specifically defined. The bitwise operators are defined assuming two’s-complement representation of integers. Numbering and counting loops generally begin from 0.

3.4.1 Arithmetic Operators

+	Addition.
–	Subtraction (as a binary operator) or negation (as a unary operator).
++	Increment.
--	Decrement.
* or ×	Multiplication.
^	Power.
/	Integer division with truncation of the result toward 0. For example, $7/4$ and $-7/-4$ are truncated to 1 and $-7/4$ and $7/-4$ are truncated to -1 .
//	Integer division with rounding to the nearest integer. Half-integer values are rounded away from 0 unless otherwise specified. For example $3//2$ is rounded to 2, and $-3//2$ is rounded to -2 .
DIV	Integer division with truncation of the result towards $-\infty$.
%	Modulus operator. Defined only for positive numbers.
Sign()	$\text{Sign}(x) = 1 \quad x > 0$ $= 0 \quad x == 0$ $= -1 \quad x < 0$
NINT ()	Nearest integer operator. Returns the nearest integer value to the real-valued argument. Half-integer values are rounded away from 0.
sin	Sine.
cos	Cosine.

exp	Exponential.
$\sqrt{\quad}$	Square root.
\log_{10}	Logarithm to base ten.
\log_e	Logarithm to base e.

3.4.2 Logical Operators

	Logical OR.
&&	Logical AND.
!	Logical NOT.

3.4.3 Relational Operators

>	Greater than.
\geq	Greater than or equal to.
<	Less than.
\leq	Less than or equal to.
==	Equal to.
!=	Not equal to.
max [,...]	The maximum value in the argument list.
min [,...]	The minimum value in the argument list.

3.4.4 Bitwise Operators

&	AND.
	OR.
>>	Shift right with sign extension.
<<	Shift left with 0 fill.

3.4.5 Assignment

=	Assignment operator.
---	----------------------

3.4.6 Mnemonics

The following mnemonics are defined to describe the different data types used in the coded bit stream.

bslbf	Bit string, left bit first, where "left" is the order in which bit strings are written in the Standard. Bit strings are written as a string of 1s and 0s within single quote marks, e.g. '1000 0001'. Blanks within a bit string are for ease of reading and have no significance.
uimbsf	Unsigned integer, most significant bit first.

The byte order of multi-byte words is most significant byte first.

3.4.7 Constants

π	3.14159265359...
-------	------------------

e 2.71828182845...

3.4.8 Method of Describing Bit Stream Syntax

Each data item in the coded bit stream described below is in bold type. It is described by its name, its length in bits, and a mnemonic for its type and order of transmission.

The action caused by a decoded data element in a bit stream depends on the value of that data element and on data elements previously decoded. The decoding of the data elements and definition of the state variables used in their decoding are described in the clauses containing the semantic description of the syntax. The following constructs are used to express the conditions when data elements are present, and are in normal type.

Note this syntax uses the “C” code convention that a variable or expression evaluating to a non-zero value is equivalent to a condition that is true.

while (condition) { data_element ... }	If the condition is true, then the group of data elements occurs next in the data stream. This repeats until the condition is not true.
do { data_element ... } while (condition)	The data element always occurs at least once. The data element is repeated until the condition is not true.
if (condition) { data_element ... }	If the condition is true, then the first group of data elements occurs next in the data stream.
else { data_element ... }	If the condition is not true, then the second group of data elements occurs next in the data stream.
for (i = 0; i < n; i++) { data_element ... }	The group of data elements occurs n times. Conditional constructs within the group of data elements may depend on the value of the loop control variable i , which is set to zero for the first occurrence, incremented to 1 for the second occurrence, and so forth.

As noted, the group of data elements may contain nested conditional constructs. For compactness, the {} are omitted when only one data element follows.

data_element []	data_element [] is an array of data. The number of data elements is indicated by the context.
data_element [n]	data_element [n] is the n+1th element of an array of data.
data_element [m] [n]	data_element [m] [n] is the m+1,n+1 th element of a two-dimensional array of data.
data_element [l] [m] [n]	data_element [l] [m] [n] is the l+1,m+1,n+1 th element of a three-dimensional array of data.
data_element [m..n]	data_element [m..n] is the inclusive range of bits between bit m and bit n in the data_element.

Decoders must include a means to look for start codes and sync bytes (transport stream) in order to begin decoding correctly, and to identify errors, erasures or insertions while decoding. The methods to identify these situations, and the actions to be taken, are not standardized.

3.4.8.1 Definition of bytealigned function

The function bytealigned() returns 1 if the current position is on a byte boundary; that is, the next bit in the bit stream is the first bit in a byte. Otherwise it returns 0.

3.4.8.2 Definition of nextbits function

The function nextbits() permits comparison of a bit string with the next bits to be decoded in the bit stream.

3.4.8.3 Definition of next_start_code function

The next_start_code() function removes any zero bit and zero byte stuffing and locates the next start code.

This function checks whether the current position is byte-aligned. If it is not, 0 stuffing bits are present. After that any number of 0 bytes may be present before the start-code. Therefore start-codes are always byte-aligned and may be preceded by any number of 0 stuffing bits.

Table 3.1 Next Start Code

Syntax	No. of bits	Mnemonic
next_start_code({		
while (!bytealigned())		
zero_bit	1	'0'
while (nextbits()!='0000 0000 0000 0000 0000 0001')		
zero_byte	8	'00000000'
}		

4. BACKGROUND

The Advanced Television Systems Committee, chaired by James C. McKinney, was formed by the member organizations of the Joint Committee on InterSociety Coordination (JCIC)¹ for the purpose of exploring the need for and, where appropriate, to coordinate development of the documentation of Advanced Television Systems. Documentation is understood to include voluntary technical standards, recommended practices, and engineering guidelines.

Proposed documentation may be developed by the ATSC, by member organizations of the JCIC, or by existing standards committees. The ATSC was established recognizing that the prompt, efficient and effective development of a coordinated set of national standards is essential to the future development of domestic television services.

On June 5, 1992, ATSC provided information to the Federal Communications Commission (FCC) outlining proposed industry actions to fully document the advanced television system standard. The FCC has recognized the importance of prompt disclosure of the system technical specifications to the mass production of advanced television system professional and consumer equipment in a timely fashion. The FCC has further noted its appreciation of the diligence with which the ATSC and the other groups participating in the standardization are pursuing these matters.²

Supporting this activity, the ATSC Executive Committee requested that the T3/S1 Specialist Group on Macro Systems Approach meet and suggest which portions of an advanced television system broadcasting standard might require action by the FCC and which portions should be voluntary.

Subsequently, T3/S1 held meetings and developed recommendations in two areas:

- 1) Principles upon which documentation of the advanced television system should be based
- 2) A list of characteristics of an advanced television system that should be documented

The list tentatively identified the industry group(s) that would provide the documentation information and the document where the information would likely appear.

The recommendations developed by the T3/S1 Specialist Group were modified by T3 to accommodate information and knowledge about advanced television systems developed in the period since June 1992. Some of the modifications to the recommendations ensued from the formation of the Grand Alliance. The modified guidelines were approved at the March 31, 1994, meeting of the T3 Technology Group on Distribution and are described in Section 4.4.

4.1 Advisory Committee on Advanced Television Service (ACATS)

A "Petition for Notice of Inquiry" was filed with the FCC on February 21, 1987, by 58 broadcasting organizations and companies requesting that the Commission initiate a proceeding

¹ The JCIC is presently composed of: the Electronic Industries Alliance (EIA), the Institute of Electrical and Electronics Engineers (IEEE), the National Association of Broadcasters (NAB), the National Cable Television Association (NCTA), and the Society of Motion Picture and Television Engineers (SMPTE).

² FCC 92-438, MM Docket No. 87-268, "Memorandum Opinion and Order/Third Report and Order/Third Further Notice of Proposed Rule Making," Adopted: September 17, 1992, pp. 59–60.

to explore the issues arising from the introduction of advanced television technologies and their possible impact on the television broadcasting service. At that time, it was generally believed that high-definition television (HDTV) could not be broadcast using 6 MHz terrestrial broadcasting channels. The broadcasting organizations were concerned that alternative media would be able to deliver HDTV to the viewing public, placing terrestrial broadcasting at a severe disadvantage.

The FCC agreed that this was a subject of utmost importance and initiated a proceeding (MM Docket No. 87-268) to consider the technical and public policy issues of advanced television systems. The Advisory Committee on Advanced Television Service was empaneled by the Federal Communications Commission in 1987, with Richard E. Wiley as chairman, to develop information that would assist the FCC in establishing an advanced television standard for the United States. The objective given to the Advisory Committee in its Charter by the FCC was:

“The Committee will advise the Federal Communications Commission on the facts and circumstances regarding advanced television systems for Commission consideration of technical and public policy issues. In the event that the Commission decides that adoption of some form of advanced broadcast television is in the public interest, the Committee would also recommend policies, standards, and regulations that would facilitate the orderly and timely introduction of advanced television services in the United States.”

The Advisory Committee established a series of subgroups to study the various issues concerning services, technical parameters, and testing mechanisms required to establish an advanced television system standard. The Advisory Committee also established a system evaluation, test, and analysis process that began with over twenty proposed systems, reducing them to four final systems for consideration.

4.2 Digital HDTV Grand Alliance (Grand Alliance)

On May 24, 1993, the three groups that had developed the four final digital systems agreed to produce a single, best-of-the best system to propose as the standard. The three groups (AT&T and Zenith Electronics Corporation; General Instrument Corporation and the Massachusetts Institute of Technology; and Philips Consumer Electronics, Thomson Consumer Electronics, and the David Sarnoff Research Center) have been working together as the “Digital HDTV Grand Alliance.” The system described in this Standard is based on the Digital HDTV Grand Alliance proposal to the Advisory Committee.

4.3 Organization for Documenting the Digital Television Standard

The ATSC Executive Committee assigned the work of documenting the advanced television system standards to T3 specialist groups, dividing the work into five areas of interest:

- Video, including input signal format and source coding
- Audio, including input signal format and source coding
- Transport, including data multiplex and channel coding
- RF/Transmission, including the modulation subsystem
- Receiver characteristics

A steering committee consisting of the chairs of the five specialist groups, the chair and vice-chairs of T3, and liaison among the ATSC, the FCC, and ACATS was established to coordinate the development of the documents. The members of the steering committee and areas of interest were as follows:

Stanley Baron	T3 chair
Jules Cohen	T3 vice-chair
Brian James	T3 vice-chair
Larry Pearlstein	T3/S6 (video systems characteristics), chair
Graham S. Stubbs	T3/S7 (audio systems characteristics), chair
Bernard J. Lechner	T3/S8 (service multiplex and transport systems characteristics), chair
Lynn D. Claudy	T3/S9 (RF/transmission systems characteristics), chair
Werner F. Wedam	T3/S10 (receiver characteristics), chair
Robert M. Rast	Grand Alliance facilitator
Robert Hopkins	ATSC
Robert M. Bromery	FCC Office of Engineering and Technology
Gordon Godfrey	FCC Mass Media Bureau
Paul E. Misener	ACATS

4.4 Principles for Documenting the Digital Television Standard

T3 adopted the following principles for documenting the advanced television system standard:

- 1) The Grand Alliance was recognized as the principal supplier of information for documenting the advanced television system, supported by the ATSC and others. Other organizations seen as suppliers of information were EIA, FCC, IEEE, MPEG, NCTA, and SMPTE.
- 2) The Grand Alliance was encouraged to begin drafting the essential elements of system details as soon as possible to avoid delays in producing the advanced television system documentation.
- 3) FCC requirements for the advanced television system standard were to be obtained as soon as possible.
- 4) Complete functional system details (permitting those skilled in the art to construct a working system) were to be made publicly available.
- 5) Protection of any intellectual property made public must be by patent or copyright as appropriate.
- 6) The advanced television system documentation shall include the necessary system information such that audio and video encoders may be manufactured to deliver the system's full demonstrated performance quality.
- 7) The advanced television system documentation shall point to existing standards, recommended practices, or guideline documents. These documents shall be referenced in one of two ways as deemed appropriate for the application. In the first instance, a specific revision shall be specified where review of changes to the referenced document is required before changes might be incorporated into the advanced television system

document. The second instance references the document without specificity to revision and allows any changes to the referenced documents to be automatically incorporated.

- 8) System specifications shall explain how future, compatible improvements may be achieved.
- 9) As ongoing improvements take place in the advanced television system, manufacturers of encoders and decoders should coordinate their efforts to insure compatibility.
- 10) The advanced television system standard must support backward compatibility of future improvements with all generations of advanced television system receivers and inherently support production of low cost receivers (notwithstanding that cost reduction through reduced performance quality may also be used to achieve inexpensive products).
- 11) The advanced television system standard should not foreclose flexibility in implementing advanced television system receivers at different price and performance levels.
- 12) The advanced television system standard should not foreclose flexibility in implementing program services or in data stream modification or insertion of data packets by downstream (local) service providers.
- 13) The advanced television system documentation shall address interoperability with non-broadcast delivery systems including cable.
- 14) The advanced television system standard shall identify critical system parameters and shall provide information as to the range of acceptable values, the method of measurement, and the location in the system where measurement takes place.

5. SYSTEM OVERVIEW

The Digital Television Standard describes a system designed to transmit high quality video and audio and ancillary data over a single 6 MHz channel. The system can deliver reliably about 19 Mbps of throughput in a 6 MHz terrestrial broadcasting channel and about 38 Mbps of throughput in a 6 MHz cable television channel. This means that encoding a video source whose resolution can be as high as five times that of conventional television (NTSC) resolution requires a bit rate reduction by a factor of 50 or higher. To achieve this bit rate reduction, the system is designed to be efficient in utilizing available channel capacity by exploiting complex video and audio compression technology.

The objective is to maximize the information passed through the data channel by minimizing the amount of data required to represent the video image sequence and its associated audio. The objective is to represent the video, audio, and data sources with as few bits as possible while preserving the level of quality required for the given application.

Although the RF/transmission subsystems described in this Standard are designed specifically for terrestrial and cable applications, the objective is that the video, audio, and service multiplex/transport subsystems be useful in other applications.

5.1 System Block Diagram

A basic block diagram representation of the system is shown in Figure 5.1. This representation is based on one adopted by the International Telecommunication Union, Radiocommunication

Sector (ITU-R), Task Group 11/3 (Digital Terrestrial Television Broadcasting). According to this model, the digital television system can be seen to consist of three subsystems.³

- Source coding and compression
- Service multiplex and transport
- RF/transmission

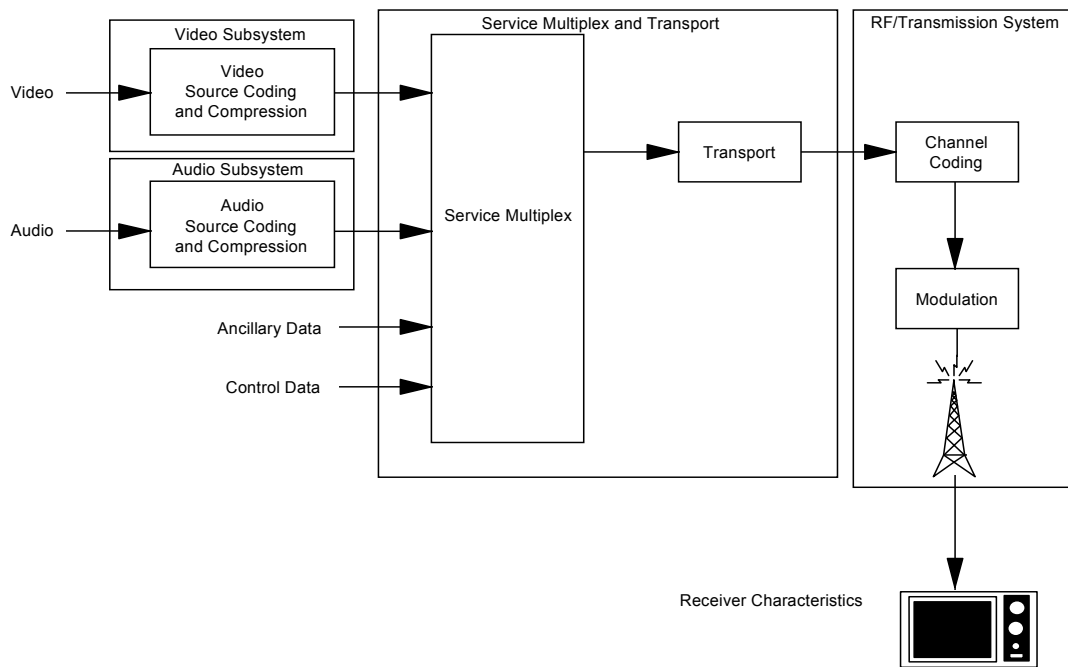


Figure 5.1 ITU-R digital terrestrial television broadcasting model.

“Source coding and compression” refers to the bit rate reduction methods, also known as data compression, appropriate for application to the video, audio, and ancillary digital data streams. The term “ancillary data” includes control data, conditional access control data, and data associated with the program audio and video services, such as closed captioning. “Ancillary data” can also refer to independent program services. The purpose of the coder is to minimize the number of bits needed to represent the audio and video information. The digital television system employs the MPEG-2 video stream syntax for the coding of video and the Digital Audio Compression (AC-3) Standard for the coding of audio.

“Service multiplex and transport” refers to the means of dividing the digital data stream into “packets” of information, the means of uniquely identifying each packet or packet type, and the appropriate methods of multiplexing video data stream packets, audio data stream packets, and ancillary data stream packets into a single data stream. In developing the transport mechanism, interoperability among digital media, such as terrestrial broadcasting, cable distribution, satellite distribution, recording media, and computer interfaces, was a prime consideration. The digital

³ ITU-R Document TG11/3-2, “Outline of Work for Task Group 11/3, Digital Terrestrial Television Broadcasting,” June 30, 1992.

television system employs the MPEG-2 transport stream syntax for the packetization and multiplexing of video, audio, and data signals for digital broadcasting systems.⁴ The MPEG-2 transport stream syntax was developed for applications where channel bandwidth or recording media capacity is limited and the requirement for an efficient transport mechanism is paramount. It was designed also to facilitate interoperability with the ATM transport mechanism.

“RF/transmission” refers to channel coding and modulation. The channel coder takes the data bit stream and adds additional information that can be used by the receiver to reconstruct the data from the received signal which, due to transmission impairments, may not accurately represent the transmitted signal. The modulation (or physical layer) uses the digital data stream information to modulate the transmitted signal. The modulation subsystem offers two modes: a terrestrial broadcast mode (8 VSB), and a high data rate mode (16 VSB).

Figure 5.2 illustrates a high level view of encoding equipment. This view is not intended to be complete, but is used to illustrate the relationship of various clock frequencies within the encoder. There are two domains within the encoder where a set of frequencies are related, the source coding domain and the channel coding domain.

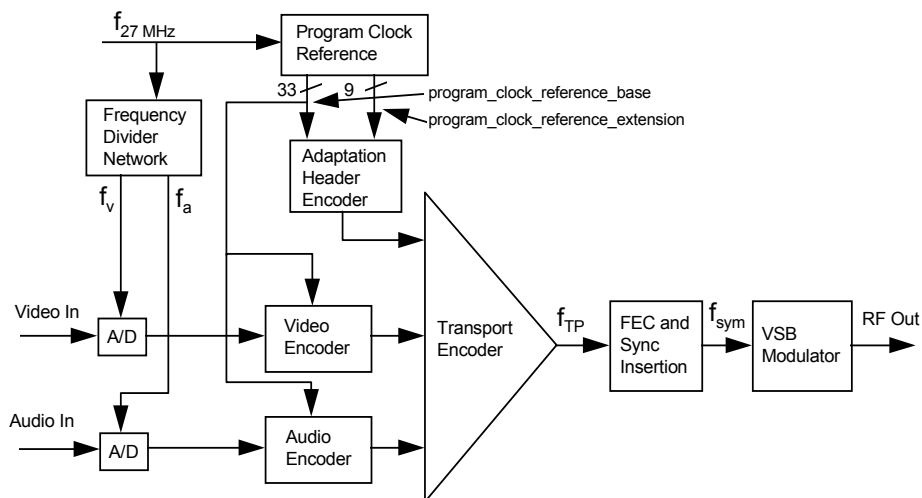


Figure 5.2 High level view of encoding equipment.

The source coding domain, represented schematically by the video, audio, and transport encoders, uses a family of frequencies which are based on a 27 MHz clock ($f_{27\text{MHz}}$). This clock is used to generate a 42-bit sample of the frequency which is partitioned into two parts defined by the MPEG-2 specification. These are the 33-bit `program_clock_reference_base` and the 9-bit `program_clock_reference_extension`. The former is equivalent to a sample of a 90 kHz clock which is locked in frequency to the 27 MHz clock, and is used by the audio and video source encoders when encoding the presentation time stamp (PTS) and the decode time stamp (DTS). The audio and video sampling clocks, f_a and f_v respectively, must be frequency-locked to the 27 MHz clock. This can be expressed as the requirement that there exist two pairs of integers, (n_a, m_a) and (n_v, m_v) , such that:

⁴ Chairman, ITU-R Task Group 11/3, “Report of the Second Meeting of ITU-R Task Group 11/3, Geneva, October 13–19, 1993,” p. 40, January 5, 1994.

$$f_a = \frac{n_a}{m_a} \times 27 \text{ MHz}$$

and

$$f_v = \frac{n_v}{m_v} \times 27 \text{ MHz}$$

The channel coding domain is represented by the FEC/Sync Insertion subsystem and the VSB modulator. The relevant frequencies in this domain are the VSB symbol frequency (f_{sym}) and the frequency of the transport stream (f_{tp}) which is the frequency of transmission of the encoded transport stream. These two frequencies must be locked, having the relation:

$$f_{tp} = 2 \times \frac{188}{208} \frac{312}{313} f_{sym}$$

The signals in the two domains are not required to be frequency-locked to each other, and in many implementations will operate asynchronously. In such systems, the frequency drift can necessitate the occasional insertion or deletion of a null packet from within the transport stream, thereby accommodating the frequency disparity.

The annexes that follow consider the characteristics of the subsystems necessary to accommodate the services envisioned.

Annex A: Video System Characteristics (Normative)

1. SCOPE

This Annex describes the characteristics of the video subsystem of the Digital Television Standard. The input formats and bit stream characteristics are described in separate sections.

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.

ISO/IEC IS 13818-1, International Standard (1996), *MPEG-2 Systems*.

ISO/IEC 13818-1: 1996/Cor. 1: 1997 (E) *Technical Corrigendum 1*.

ISO/IEC 13818-1: 1996/Amd. 1: 1997 (E) *Amendment 1*.

ISO/IEC 13818-1: 1996/Amd. 2: 1997 (E) *Amendment 2*.

ISO/IEC 13818-1: 1996/Amd. 3: 1997 (E) *Amendment 3*.

ISO/IEC 13818-1: 1996/Amd. 4: 1997 (E) *Amendment 4*.

ISO/IEC IS 13818-2, International Standard (1996), *MPEG-2 Video*.

ISO/IEC 13818-2: 1996/Cor. 1: 1997 (E) *MPEG-2 Video Technical Corrigendum 1*.

ISO/IEC 13818-2: 1996/Cor. 2: 1997 (E) *MPEG-2 Video Technical Corrigendum 2*.

2.2 Informative References

SMPTE 274M (1995), Standard for Television, 1920 × 1080 Scanning and Interface, Society of Motion Picture and Television Engineers, White Plains, N.Y.

SMPTE 296M (1997), Standard for Television, 1280 × 720 Scanning, Analog and Digital Representation, and Analog Interface, Society of Motion Picture and Television Engineers, White Plains, N.Y.

ITU-R BT.601-4 (1994), Encoding parameters of digital television for studios.

3. COMPLIANCE NOTATION

As used in this document, “shall” or “will” 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 A1, that define video formats that relate to compression formats specified by this standard.

Table A1 Standardized Video Input Formats

Video Standard	Active Lines	Active Samples/ Line
SMPTE 274M	1080	1920
SMPTE 296M	720	1280
ITU-R BT.601-4	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 ATV video compression algorithm shall conform to the Main Profile syntax of ISO/IEC 13818-2, including Technical Corrigendum 1 and Technical Corrigendum 2. The allowable parameters shall be bounded by the upper limits specified for the Main Profile at High Level.¹ Additionally, ATV bit streams shall meet the constraints and specifications described in Sections 5.1 and 5.2.

5.1 Constraints with Respect to ISO/IEC 13818-2 Main Profile

The following tables list the allowed values for each of the ISO/IEC 13818-2 syntactic elements which are restricted beyond the limits imposed by MP@HL.

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.

5.1.1 Sequence Header Constraints

Table A2 identifies parameters in the sequence header of a bit stream that shall be constrained by the video subsystem and lists the allowed values for each.

Table A2 Sequence Header Constraints

Sequence Header Syntactic Element	Allowed Value
horizontal_size_value	see Table A3
vertical_size_value	see Table A3
aspect_ratio_information	see Table A3
frame_rate_code	see Table A3
bit_rate_value (≤ 19.4 Mbps)	≤ 48500
bit_rate_value (≤ 38.8 Mbps)	≤ 97000
vbv_buffer_size_value	≤ 488

The allowable values for the field `bit_rate_value` are application-dependent. In the primary application of terrestrial broadcast, this field shall correspond to a bit rate which is less than or

¹ See ISO/IEC 13818-2, Section 8 for more information regarding profiles and levels.

equal to 19.4 Mbps. In the high data rate mode, the corresponding bit rate is less than or equal to 38.8 Mbps.

5.1.2 Compression Format Constraints

Table A3 lists the allowed compression formats.

Table A3 Compression Format Constraints

vertical_size_value	horizontal_size_value	aspect_ratio_information	frame_rate_code	progressive_sequence
1080 ²	1920	1,3	1,2,4,5	1
			4,5	0
720	1280	1,3	1,2,4,5,7,8	1
480	704	2,3	1,2,4,5,7,8	1
			4,5	0
	640	1,2	1,2,4,5,7,8	1
			4,5	0
Legend for MPEG-2 coded values:				
aspect_ratio_information: 1 = square samples, 2 = 4:3 display aspect ratio, 3 = 16:9 display aspect ratio				
frame_rate_code: 1 = 23.976 Hz, 2 = 24 Hz, 4 = 29.97 Hz, 5 = 30 Hz, 7 = 59.94 Hz, 8 = 60 Hz				
progressive_sequence: 0 = interlaced scan, 1 = progressive scan				

5.1.3 Sequence Extension Constraints

Table A4 identifies parameters in the sequence extension part of a bit stream that shall be constrained by the video subsystem and lists the allowed values for each. A sequence_extension structure is required to be present after every sequence_header structure.

Table A4 Sequence Extension Constraints

Sequence Extension Syntactic Element	Allowed Values
progressive_sequence	see Table A3
profile_and_level_indication	see Note
chroma_format	'01'
horizontal_size_extension	'00'
vertical_size_extension	'00'
bit_rate_extension	'0000 0000 0000'
vbv_buffer_size_extension	'0000 0000'
frame_rate_extension_n	'00'
frame_rate_extension_d	'0000 0'

² Note that 1088 lines are actually coded in order to satisfy the MPEG-2 requirement that the coded vertical size be a multiple of 16 (progressive scan) or 32 (interlaced scan). The 1080 lines are required to start at the first line of the field; i.e., top-aligned.

Note: The `profile_and_level_indication` field shall indicate the lowest profile and level defined in ISO/IEC 13818-2, Section 8, that is consistent with the parameters of the video elementary stream.

5.1.4 Sequence Display Extension Constraints

Table A5 identifies parameters in the sequence display extension part of a bit stream that shall be constrained by the video subsystem and lists the allowed values for each.

Table A5 Sequence Display Extension Constraints

Sequence display extension syntactic element	Allowed values
<code>video_format</code>	'000'

The values for `color primaries`, `transfer characteristics`, and `matrix coefficients` shall be explicitly indicated in the `sequence_display_extension`. While all values for `color primaries`, `transfer characteristics`, and `matrix coefficients` defined in Tables 6-7, 6-8, and 6-9 of ISO/IEC 13818-2 are allowed in the transmitted bit stream, it is noted that ITU-R BT.709 and SMPTE 170M are the most likely to be in common use.

Note: Some previously-encoded legacy material may not have the colorimetry (i.e., `color primaries`, `transfer characteristics`, and `matrix coefficients`) explicitly indicated in the `sequence_display_extension`, in which case the colorimetry is most likely ITU-R BT.709 for all formats except those formats with `vertical_size_value` = 480, which are most likely to have colorimetry according to SMPTE 170M.

5.1.5 Picture Header Constraints

In all cases other than when `vbv_delay` has the value 0xFFFF, the value of `vbv_delay` shall be constrained as follows³:

$$\text{vbv_delay} \leq 45000$$

5.1.6 Picture Coding Constraints

`frame_pred_frame_dct` shall be '1' if `progressive_frame` is '1'

5.2 Bit Stream Specifications Beyond MPEG-2

This section covers the extension and user data part of the video syntax. These data are inserted at the sequence, GOP, and picture level. The syntax used for the insertion of closed captioning in picture user data is described.⁴

5.2.1 Picture Extension and User Data Syntax

Table A6 describes the syntax used for picture extension and user data.

³ Note: This publication corrects a typographical error that incorrectly indicated “`vbv_delay` = 45000”.

⁴ In order to decode the user data, the decoder should properly recognize the 32-bit ATSC registration identifier at the PSI stream level (see ISO/IEC 13818-1).

Table A6 Picture Extension and User Data Syntax

Value	No. of Bits	Mnemonic
extension_and_user_data(2) {		
while ((nextbits() == extension_start_code) (nextbits() == user_data_start_code)) {		
if (nextbits() == extension_start_code)		
extension_data(2)		
if (nextbits() == user_data_start_code)		
user_data(2)		
}		
}		

5.2.2 Picture User Data Syntax

Table A7 describes the picture user data syntax. (See next page.)

Table A7 Picture User Data Syntax⁵

Value	No. of Bits	Mnemonic
<code>user_data() {</code>		
<code>user_data_start_code</code>	32	bslbf
<code>ATSC_identifier</code>	32	bslbf
<code>user_data_type_code</code>	8	uimsbf
<code>if (user_data_type_code == '0x03') {</code>		
<code>process_em_data_flag</code>	1	bslbf
<code>process_cc_data_flag</code>	1	bslbf
<code>additional_data_flag</code>	1	bslbf
<code>cc_count</code>	5	uimsbf
<code>em_data</code>	8	bslbf
<code>for (i=0 ; i < cc_count ; i++) {</code>		
<code>marker_bits</code>	5	'1111 1'
<code>cc_valid</code>	1	bslbf
<code>cc_type</code>	2	bslbf
<code>cc_data_1</code>	8	bslbf
<code>cc_data_2</code>	8	bslbf
<code>}</code>		
<code>marker_bits</code>	8	'1111 1111'
<code>if (additional_data_flag) {</code>		
<code>while (nextbits() != '0000 0000 0000 0000 0000 0001') {</code>		
<code>additional_user_data</code>	8	
<code>}</code>		
<code>}</code>		
<code>} else {</code>		
<code>while (nextbits() != '0000 0000 0000 0000 0000 0001') {</code>		
<code>ATSC_reserved_user_data</code>	8	
<code>}</code>		
<code>}</code>		
<code>next_start_code()</code>		
<code>}</code>		

5.2.3 Picture User Data Semantics

user_data_start_code: This is set to 0x0000 01B2.

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.

⁵ Shaded cells in this table indicate syntactic and semantic additions to the ISO/IEC 13818-2 standard.

user_data_type_code: The 8-bit code is set to 0x03 for EIA advanced television closed captioning. All other values are reserved for use by ATSC.

process_em_data_flag: This flag is set to indicate whether it is necessary to process the `em_data`. If it is set to 1, the `em_data` has to be parsed and its meaning has to be processed. When it is set to 0, the `em_data` can be discarded.

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`.

em_data: Eight bits for representing emergency message.⁶

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.

cc_type: Denotes the type of the two closed caption data bytes that follow.⁷

cc_data_1: The first byte of a closed caption data pair.

cc_data_2: The second byte of a closed caption data pair.

additional_user_data: Any further demand for picture user data could be met by defining this part of the bit stream.

ATSC_reserved_user_data: Reserved for future use by ATSC.

⁶ There is currently no activity underway to define the interpretation of this field.

⁷ EIA, *Recommended Practice for Advanced Television Closed Captioning*, draft, July 1, 1994.

Annex B: Audio System Characteristics (Normative)

1. SCOPE

This Annex describes the audio system characteristics and normative specifications of the Digital Television Standard.

2. NORMATIVE REFERENCES

The following documents contain provisions which in whole or part, 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 amendment, 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.

ATSC Standard A/52A (2001), Digital Audio Compression (AC-3).

AES 3-1992 (ANSI S4.40-1992), AES Recommended Practice for digital audio engineering—Serial transmission format for two-channel linearly represented digital audio data.

ANSI S1.4-1983, Specification for Sound Level Meters.

IEC 651 (1979), Sound Level Meters.

IEC 804 (1985), Amendment 1 (1989) Integrating/Averaging Sound Level Meters.

3. COMPLIANCE NOTATION

As used in this document, “shall” or “will” 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. SYSTEM OVERVIEW

As illustrated in Figure B1, the audio subsystem comprises the audio encoding/decoding function and resides between the audio inputs/outputs and the transport subsystem. The audio encoder(s) is (are) responsible for generating the audio elementary stream(s) which are encoded representations of the baseband audio input signals. At the receiver, the audio subsystem is responsible for decoding the audio elementary stream(s) back into baseband audio.

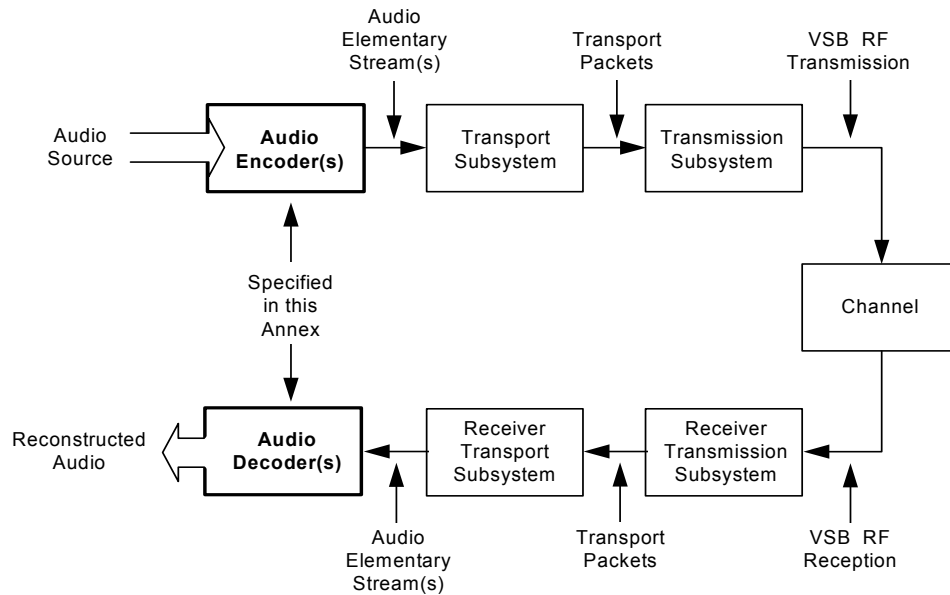


Figure B1 Audio subsystem in the digital television system.

5. SPECIFICATION

This Section forms the normative specification of the audio system. The audio compression system conforms with the Digital Audio Compression (AC-3) Standard, subject to the constraints outlined in this Section.

5.1 Constraints With Respect to ATSC Standard A/52A

The digital television audio coding system is based on the Digital Audio Compression (AC-3) Standard specified in the body of ATSC Doc. A/52A (the non-normative annexes are not included). Constraints on the system are shown in Table B1, which shows permitted values of certain syntactical elements. These constraints are described in Sections 5.2–5.4. The receive audio buffer is specified in Section 5.7.

Table B1 Audio Constraints

AC-3 Syntactical Element	Comment	Allowed value
fscod	Indicates sampling rate	'00' (indicates 48 kHz)
frmsizecod	Main audio service or associated audio service containing all necessary program elements	≤ '011110' (indicates ≤ 448 kb/s)
frmsizecod	Single channel associated service containing a single program element	≤ '010000' (indicates ≤ 128 kbps)
frmsizecod	Two channel dialogue associated service	≤ '010100' (indicates ≤ 192 kbps)
(frmsizecod)	Combined bit rate of a main and an associated service intended to be simultaneously decoded	(total ≤ 576 kbps)
acmod	Indicates number of channels	≥ '001'

5.2 Sampling Frequency

The system conveys digital audio sampled at a frequency of 48 kHz, locked to the 27 MHz system clock. The 48 kHz audio sampling clock is defined as:

$$48 \text{ kHz audio sample rate} = (2 \div 1125) \times (27 \text{ MHz system clock})$$

If analog signal inputs are employed, the A/D converters should sample at 48 kHz. If digital inputs are employed, the input sampling rate shall be 48 kHz, or the audio encoder shall contain sampling rate converters which convert the sampling rate to 48 kHz.

5.3 Bit Rate

A main audio service, or an associated audio service which is a complete service (containing all necessary program elements) shall be encoded at a bit rate less than or equal to 448 kbps. A single channel associated service containing a single program element shall be encoded at a bit rate less than or equal to 128 kbps. A two channel associated service containing only dialogue shall be encoded at a bit rate less than or equal to 192 kbps. The combined bit rate of a main service and an associated service which are intended to be decoded simultaneously shall be less than or equal to 576 kbps.

5.4 Audio Coding Modes

Audio services shall be encoded using any of the audio coding modes specified in A/52, with the exception of the 1+1 mode. The value of *acmod* in the AC-3 bit stream shall have a value in the range of 1–7, with the value 0 prohibited.

5.5 Dialogue Level

The value of the *dialnorm* parameter in the AC-3 elementary bit stream shall indicate the level of average spoken dialogue within the encoded audio program. Dialogue level may be measured by means of an “A” weighted integrated measurement (*L_{Aeq}*). (Receivers use the value of *dialnorm* to adjust the reproduced audio level so as to normalize the dialogue level.)

5.6 Dynamic Range Compression

Each encoded audio block may contain a dynamic range control word (*dynrng*) that is used by decoders (by default) to alter the level of the reproduced audio. The control words allow the decoded signal level to be increased or decreased by up to 24 dB. In general, elementary streams may have dynamic range control words inserted or modified without affecting the encoded audio. When it is necessary to alter the dynamic range of audio programs which are broadcast, the dynamic range control word should be used.

5.7 STD Audio Buffer Size

The main audio buffer (*BS_n*, see A/52 Annex A) shall be 2592 bytes.

6. MAIN AND ASSOCIATED SERVICES

An AC-3 elementary stream contains the encoded representation of a single audio service. Multiple audio services are provided by multiple elementary streams. Each elementary stream is conveyed by the transport multiplex with a unique PID. There are a number of audio service types

which may (individually) be coded into each elementary stream. Each AC-3 elementary stream is tagged as to its service type using the *bsmod* bit field. There are two types of *main service* and six types of *associated service*. Each associated service may be tagged (in the AC-3 audio descriptor in the transport PSI data) as being associated with one or more main audio services. Each AC-3 elementary stream may also be tagged with a language code.

Associated services may contain complete program mixes, or may contain only a single program element. Associated services which are complete mixes may be decoded and used as is. They are identified by the *full_svc* bit in the AC-3 descriptor (see A/52A, Annex A). Associated services which contain only a single program element are intended to be combined with the program elements from a main audio service.

This section specifies the meaning and use of each type of service. In general, a complete audio program (what is presented to the listener over the set of loudspeakers) may consist of a main audio service, an associated audio service that is a complete mix, or a main audio service combined with an associated audio service. The capability to simultaneously decode one main service and one associated service is required in order to form a complete audio program in certain service combinations described in this section. This capability may not exist in some receivers.

6.1 Summary of Service Types

The audio service types are listed in Table B2.

Table B2 Audio Service Types

bsmod	Type of Service
000 (0)	Main audio service: complete main (CM)
001 (1)	Main audio service: music and effects (ME)
010 (2)	Associated service: visually impaired (VI)
011 (3)	Associated service: hearing impaired (HI)
100 (4)	Associated service: dialogue (D)
101 (5)	Associated service: commentary (C)
110 (6)	Associated service: emergency (E)
111 (7)	Associated service: voice-over (VO)

6.2 Complete Main Audio Service (CM)

The CM type of main audio service contains a complete audio program (complete with dialogue, music, and effects). This is the type of audio service normally provided. The CM service may contain from 1 to 5.1 audio channels. The CM service may be further enhanced by means of the VI, HI, C, E, or VO associated services described below. Audio in multiple languages may be provided by supplying multiple CM services, each in a different language.

6.3 Main Audio Service, Music and Effects (ME)

The ME type of main audio service contains the music and effects of an audio program, but not the dialogue for the program. The ME service may contain from 1 to 5.1 audio channels. The primary program dialogue is missing and (if any exists) is supplied by simultaneously encoding

a D associated service. Multiple D associated services in different languages may be associated with a single ME service.

6.4 Visually Impaired (VI)

The VI associated service typically contains a narrative description of the visual program content. In this case, the VI service shall be a single audio channel. The simultaneous reproduction of both the VI associated service and the CM main audio service allows the visually impaired user to enjoy the main multi-channel audio program, as well as to follow (by ear) the on-screen activity.

The dynamic range control signal in this type of VI service is intended to be used by the audio decoder to modify the level of the main audio program. Thus the level of the main audio service will be under the control of the VI service provider, and the provider may signal the decoder (by altering the dynamic range control words embedded in the VI audio elementary stream) to reduce the level of the main audio service by up to 24 dB in order to assure that the narrative description is intelligible.

Besides providing the VI service as a single narrative channel, the VI service may be provided as a complete program mix containing music, effects, dialogue, and the narration. In this case, the service may be coded using any number of channels (up to 5.1), and the dynamic range control signal applies only to this service. The fact that the service is a complete mix shall be indicated in the AC-3 descriptor (see A/52A, Annex A).

6.5 Hearing Impaired (HI)

The HI associated service typically contains only dialogue which is intended to be reproduced simultaneously with the CM service. In this case, the HI service shall be a single audio channel. This dialogue may have been processed for improved intelligibility by hearing impaired listeners. Simultaneous reproduction of both the CM and HI services allows the hearing impaired listener to hear a mix of the CM and HI services in order to emphasize the dialogue while still providing some music and effects.

Besides providing the HI service as a single dialogue channel, the HI service may be provided as a complete program mix containing music, effects, and dialogue with enhanced intelligibility. In this case, the service may be coded using any number of channels (up to 5.1). The fact that the service is a complete mix shall be indicated in the AC-3 descriptor (see A/52A, Annex A).

6.6 Dialogue (D)

The D associated service contains program dialogue intended for use with an ME main audio service. The language of the D service is indicated in the AC-3 bit stream, and in the audio descriptor. A complete audio program is formed by simultaneously decoding the D service and the ME service and mixing the D service into the center channel of the ME main service (with which it is associated).

If the ME main audio service contains more than two audio channels, the D service shall be monophonic (1/0 mode). If the main audio service contains two channels, the D service may also contain two channels (2/0 mode). In this case, a complete audio program is formed by simultaneously decoding the D service and the ME service, mixing the left channel of the ME

service with the left channel of the D service, and mixing the right channel of the ME service with the right channel of the D service. The result will be a two channel stereo signal containing music, effects, and dialogue.

Audio in multiple languages may be provided by supplying multiple D services (each in a different language) along with a single ME service. This is more efficient than providing multiple CM services, but, in the case of more than two audio channels in the ME service, requires that dialogue be restricted to the center channel.

Some receivers may not have the capability to simultaneously decode an ME and a D service.

6.7 Commentary (C)

The commentary associated service is similar to the D service, except that instead of conveying essential program dialogue, the C service conveys optional program commentary. The C service may be a single audio channel containing only the commentary content. In this case, simultaneous reproduction of a C service and a CM service will allow the listener to hear the added program commentary.

The dynamic range control signal in the single channel C service is intended to be used by the audio decoder to modify the level of the main audio program. Thus the level of the main audio service will be under the control of the C service provider, and the provider may signal the decoder (by altering the dynamic range control words embedded in the C audio elementary stream) to reduce the level of the main audio service by up to 24 dB in order to assure that the commentary is intelligible.

Besides providing the C service as a single commentary channel, the C service may be provided as a complete program mix containing music, effects, dialogue, and the commentary. In this case the service may be provided using any number of channels (up to 5.1). The fact that the service is a complete mix shall be indicated in the AC-3 descriptor (see A/52A, Annex A).

6.8 Emergency (E)

The E associated service is intended to allow the insertion of emergency or high priority announcements. The E service is always a single audio channel. An E service is given priority in transport and in audio decoding. Whenever the E service is present, it will be delivered to the audio decoder. Whenever the audio decoder receives an E type associated service, it will stop reproducing any main service being received and only reproduce the E service out of the center channel (or left and right channels if a center loudspeaker does not exist). The E service may also be used for non-emergency applications. It may be used whenever the broadcaster wishes to force all decoders to quit reproducing the main audio program and reproduce a higher priority single audio channel.

6.9 Voice-Over (VO)

The VO associated service is a single channel service intended to be reproduced along with the main audio service in the receiver. It allows typical voice-overs to be added to an already encoded audio elementary stream without requiring the audio to be decoded back to baseband and then re-encoded. It is always a single audio channel. It has second priority (only the E service has higher priority). It is intended to be simultaneously decoded and mixed into the

center channel of the main audio service. The dynamic range control signal in the VO service is intended to be used by the audio decoder to modify the level of the main audio program. Thus the level of the main audio service may be controlled by the broadcaster, and the broadcaster may signal the decoder (by altering the dynamic range control words embedded in the VO audio elementary stream) to reduce the level of the main audio service by up to 24 dB during the voice-over.

Some receivers may not have the capability to simultaneously decode and reproduce a voice-over service along with a program audio service.

7. AUDIO ENCODER INTERFACES

7.1 Audio Encoder Input Characteristics

Audio signals which are input to the digital television system may be in analog or digital form. Audio signals should have any dc offset removed before being encoded. If the audio encoder does not include a dc blocking high-pass filter, the audio signals should be high-pass-filtered before being applied to the encoder. In general, input signals should be quantized to at least 16-bit resolution. The audio compression system can convey audio signals with up to 24-bit resolution. Physical interfaces for the audio inputs to the encoder may be defined as voluntary industry standards by the AES, SMPTE, or other standards organizations.

7.2 Audio Encoder Output Characteristics

Conceptually, the output of the audio encoder is an elementary stream which is formed into PES packets within the transport subsystem. It is possible that systems will be implemented wherein the formation of audio PES packets takes place within the audio encoder. In this case, the output(s) of the audio encoder(s) would be PES packets. Physical interfaces for these outputs (elementary streams and/or PES packets) may be defined as voluntary industry standards by SMPTE or other standards organizations.

Annex C: Service Multiplex and Transport System Characteristics (Normative)

1. SCOPE

This Annex describes the transport layer characteristics and normative specifications of the Digital Television Standard.

2. NORMATIVE REFERENCES

The following documents contain provisions which in whole or in part, 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 amendment, and parties to agreements based on this Standard are encouraged to investigate the possibility of applying the most recent editions of the documents listed below.

ATSC Standard A/52 (1995), Digital Audio Compression (AC-3).

ATSC Standard A/65 (1997), Program and System Information Protocol for Terrestrial Broadcast and Cable.

ISO/IEC IS 13818-1, International Standard (1996), MPEG-2 Systems.

ISO/IEC 13818-1: 1996/Cor. 1: 1997 (E) Technical Corrigendum 1.

ISO/IEC 13818-1: 1996/Amd. 1: 1997 (E) Amendment 1.

ISO/IEC 13818-1: 1996/Amd. 2: 1997 (E) Amendment 2.

ISO/IEC 13818-1: 1996/Amd. 3: 1997 (E) Amendment 3.

ISO/IEC 13818-1: 1996/Amd. 4: 1997 (E) Amendment 4.

ISO/IEC IS 13818-2, International Standard (1996), MPEG-2 Video.

ISO/IEC 13818-2: 1996/Cor. 1: 1997 (E) MPEG-2 Video Technical Corrigendum 1.

ISO/IEC 13818-2: 1996/Cor. 2: 1997 (E) MPEG-2 Video Technical Corrigendum 2.

ISO/IEC CD 13818-4, MPEG Committee Draft (1994), MPEG-2 Compliance.

3. COMPLIANCE NOTATION

As used in this document, “shall” or “will” 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. SYSTEM OVERVIEW

The transport format and protocol for the Digital Television Standard is a compatible subset of the MPEG-2 Systems specification defined in ISO/IEC 13818-1. It is based on a fixed-length packet transport stream approach which has been defined and optimized for digital television delivery applications.

As illustrated in Figure C1, the transport function resides between the application (e.g., audio or video) encoding and decoding functions and the transmission subsystem. The encoder's transport subsystem is responsible for formatting the coded elementary streams and multiplexing the different components of the program for transmission. At the receiver, it is responsible for recovering the elementary streams for the individual application decoders and for the corresponding error signaling. The transport subsystem also incorporates other higher protocol layer functionality related to synchronization of the receiver.

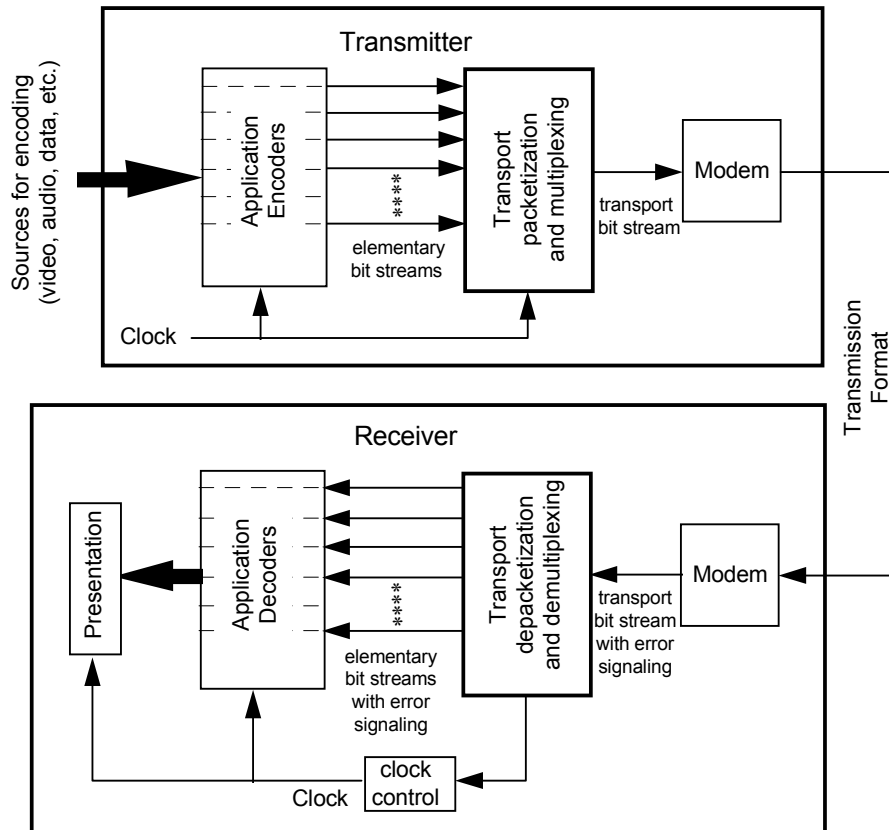


Figure C1 Sample organization of functionality in a transmitter-receiver pair for a single program.

The overall system multiplexing approach can be thought of as a combination of multiplexing at two different layers. In the first layer, single program transport bit streams are formed by multiplexing transport packets from one or more Packetized Elementary Stream (PES) sources. In the second layer, many single program transport bit streams are combined to form a system of programs. The Program Specific Information (PSI) streams contain the information relating to the identification of programs and the components of each program.

Not shown explicitly in Figure C1, but essential to the practical implementation of this Standard, is a control system that manages the transfer and processing of the elementary streams from the application encoders. The rules followed by this control system are not a part of this

Standard but must be established as recommended practices by the users of the Standard. The control system implementation shall adhere to the requirements of the MPEG-2 transport system as specified in ISO/IEC 13818-1 with the additional constraints specified in this Standard. These constraints may go beyond the constraints imposed by the application encoders.

5. SPECIFICATION

This section constitutes the normative specification for the transport system of the Digital Television Standard. The syntax and semantics of the specification conform to ISO/IEC 13818-1 subject to the constraints and conditions specified in this Standard. This section of the Standard describes the coding constraints that apply to the use of the MPEG-2 systems specification in the digital television system.

5.1 MPEG-2 Systems Standard

The transport system is based on the transport stream definition of the MPEG-2 Systems standard as specified in ISO/IEC 13818-1.

5.1.1 Video T-STD

The video T-STD is specified in Section 2.4.2.3 of ISO/IEC 13818-1 and follows the constraints for the level encoded in the video elementary stream.

5.1.2 Audio T-STD

The audio T-STD is specified in Section 3.6 of Annex A of ATSC Standard A/52.

5.2 Registration Descriptor

This Standard uses the registration descriptor described in Section 2.6.8 of ISO/IEC 13818-1 to identify the contents of programs and elementary streams to decoding equipment.

5.2.1 Program Identifier

Programs which conform to this specification will be identified by the 32-bit identifier in the section of the Program Map Table (PMT) detailed in Section 2.4.4.8 of ISO/IEC 13818-1. The identifier will be coded according to Section 2.6.8, and shall have a value of 0x4741 3934.

5.2.2 Audio Elementary Stream Identifier

Audio elementary streams which conform to this specification will be identified by the 32-bit identifier in the section of the Program Map Table (PMT) detailed in Section 2.4.4.8 of ISO/IEC 13818-1. The identifier will be coded according to Section 2.6.8, and shall have a value of 0x4143 2D33.

5.3 Audio Constraints

If a program contains an audio component, the primary audio shall be a complete main audio service (CM) as defined by ATSC Standard A/52 and shall contain the complete primary audio of the program including all required voice-overs and emergency messages.

5.4 Constraints on PSI

The program constituents for all programs are described in the PSI. There are the following constraints on the PSI information:

- Only one program is described in a PSI transport bit stream corresponding to a particular PMT_PID value. For terrestrial broadcast applications, a transport bit stream containing a program_map_table shall not be used to transmit any other kind of PSI table (identified by a different table_id).
- The maximum spacing between occurrences of a program_map_table containing television program information shall be 400 ms.
- The program numbers are associated with the corresponding PMT_PIDs in the PID0 Program Association Table. The maximum spacing between occurrences of section 0 of the program_association_table is 100 ms.
- The video elementary stream section shall contain the Data stream alignment descriptor described in Section 2.6.10 of ISO/IEC 13818-1. The alignment_type field shown in Table 2-47 of ISO/IEC 13818-1 shall be 0x02.
- Adaptation headers shall not occur in transport packets of the PMT_PID for purposes other than for signaling with the discontinuity_indicator that the version_number (Section 2.4.4.5 of ISO/IEC 13818-1) may be discontinuous.
- Adaptation headers shall not occur in transport packets of the PAT_PID for purposes other than for signaling with the discontinuity_indicator that the version_number (Section 2.4.4.5 of ISO/IEC 13818-1) may be discontinuous.

5.5 PES Constraints

Packetized Elementary Stream syntax and semantics shall be used to encapsulate the audio and video elementary stream information. The Packetized Elementary Stream syntax is used to convey the Presentation Time-Stamp (PTS) and Decoding Time-Stamp (DTS) information required for decoding audio and video information with synchronism. This section describes the coding constraints for this system layer.

Within the PES packet header, the following restrictions apply:

- PES_scrambling_control shall be coded as '00'.
- ESCR_flag shall be coded as '0'.
- ES_rate_flag shall be coded as '0'.
- PES_CRC_flag shall be coded as '0'.

Within the PES packet extension, the following restrictions apply.

- PES_private_data_flag shall be coded as '0'.
- pack_header_field_flag shall be coded as '0'.
- program_packet_sequence_counter_flag shall be coded as '0'.
- P-STD_buffer_flag shall be coded as '0'.

5.5.1 Video PES Constraints

Each PES packet shall begin with a video access unit, as defined in Section 2.1.1 of ISO/IEC 13818-1, which is aligned with the PES packet header. The first byte of a PES packet payload

shall be the first byte of a video access unit. Each PES header shall contain a PTS. Additionally, it shall contain a DTS as appropriate. For terrestrial broadcast, the PES packet shall not contain more than one coded video frame, and shall be void of video picture data only when transmitted in conjunction with the `discontinuity_indicator` to signal that the `continuity_counter` may be discontinuous.

Within the PES packet header, the following restrictions apply:

- The `PES_packet_length` shall be coded as ‘0x0000’
- `data_alignment_indicator` shall be coded as ‘1’

5.5.2 Audio PES Constraints

The audio decoder may be capable of simultaneously decoding more than one elementary stream containing different program elements, and then combining the program elements into a complete program. In this case, the audio decoder may sequentially decode audio frames (or audio blocks) from each elementary stream and do the combining (mixing together) on a frame or (block) basis. In order to have the audio from the two elementary streams reproduced in exact sample synchronism, it is necessary for the original audio elementary stream encoders to have encoded the two audio program elements frame synchronously; i.e., if audio program 1 has sample 0 of frame n at time t_0 , then audio program 2 should also have frame n beginning with its sample 0 at the identical time t_0 . If the encoding is done frame synchronously, then matching audio frames should have identical values of PTS.

If PES packets from two audio services that are to be decoded simultaneously contain identical values of PTS then the corresponding encoded audio frames contained in the PES packets should be presented to the audio decoder for simultaneous synchronous decoding. If the PTS values do not match (indicating that the audio encoding was not frame synchronous) then the audio frames which are closest in time may be presented to the audio decoder for simultaneous decoding. In this case, the two services may be reproduced out of sync by as much as 1/2 of a frame time (which is often satisfactory, e.g., a voice-over does not require precise timing).

The value of `stream_id` for AC-3 shall be 1011 1101 (`private_stream_1`).

5.6 Services and Features

5.6.1 System Information and Program Guide

5.6.1.1 System information and program guide PID

Compliant Transport Streams shall include system information and program guide data formatted according to the structure and syntax described in ATSC Standard A/65, “Program and System Information Protocol for Terrestrial Broadcast and Cable.” System information and program guide data shall be conveyed in Transport Stream packets of PID 0x1FFB, which shall be reserved exclusively for this purpose. System information provides data necessary for navigation among digital service offerings. The program guide database allows a receiver to build an on-screen grid of program information for the various services that may be available.

5.6.1.2 System information and program guide STD model

The STD model for program guide and system information is specified in ATSC Standard A/65A.

5.6.2 Specification of Private Data Services

Private data provides a means to add new ancillary services to the basic digital television service specified in this standard. Private data is supported in two bit stream locations.

- 1) Private data can be transmitted within the adaptation header of transport packets (Sections 2.4.3.4 and 2.4.3.5 of ISO/IEC 13818-1).
- 2) Private data can be transmitted as a separate transport stream with its own PID. The contents can be identified as being ATSC private by using the `private_data_indicator_descriptor` (Section 2.6.29 of ISO/IEC 13818-1) within the PMT.

In either case, it is necessary that the standards which specify the characteristics of such `private_streams` be consistent with the Digital Television Standard. Standards for `private_streams` shall precisely specify the semantics of the transmitted syntax as described in Sections 5.6.2.1 and 5.6.2.1.1.

5.6.2.1 Verification model

The standard shall be specified in terms of a verification model by defining the characteristics of the transmitted syntax and an idealized decoder. In ISO/IEC 13818-1 and 13818-2, this is accomplished by using the T-STD and VBV models, respectively. The elements required for specification by this Standard are described in the following Sections.

5.6.2.1.1 Syntax and semantics

The syntax and semantics of the transmitted bit stream that implements the ancillary service shall be completely and unambiguously specified. The decoding process shall also be completely and unambiguously specified.

5.6.2.1.2 Ancillary service target decoder (ASTD)

An idealized decoder model must be precisely defined for the service. Figure C2 introduces a concrete model for pedagogic purposes. It is modeled after the T-STD.

The salient features of the model are the size of the transport demultiplexing buffer (TB), the minimum transfer rate out of the transport demultiplex buffer (R_{leak}), the required System buffering (BS_{sys}), and optionally the partitioning of BS_{sys} between the smoothing portion and the decoder portion. The decoding process, represented as the decoding times $\tau_{\text{decode}(i)}$, must be completely specified. The behavior of the BS_{sys} buffer must be completely modeled with respect to its input process and its output process. Certain parameters of the service such as bit rate, etc., should also be specified.

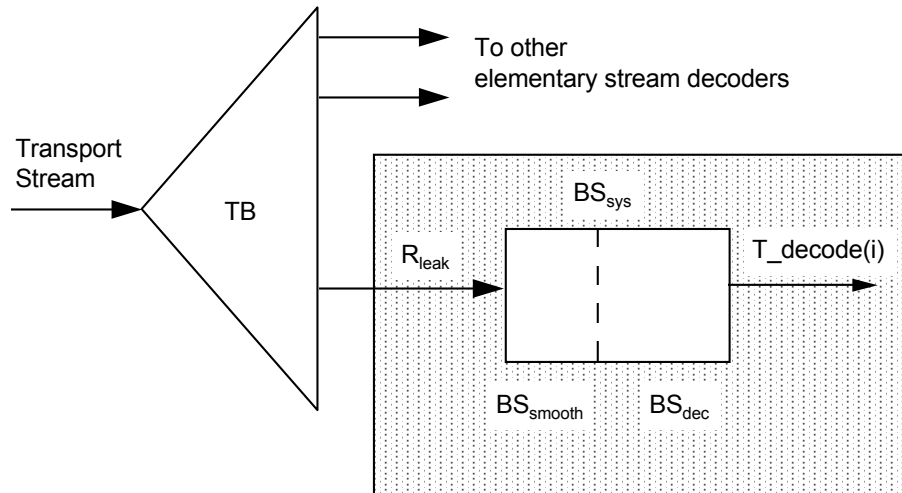


Figure C2 Ancillary service target decoder.

5.6.2.2 Stream type and PMT descriptors

A new ancillary service shall be described as a program or elementary stream through documented Program Specific Information.

5.6.2.2.1 Stream type

Several identifiers that are part of the transport section of the Digital Television Standard may be used to identify either the signal or constituent parts thereof; however, the fundamental identifier is the User Private stream type. The `stream_type` codes shall be unambiguously assigned within the range 0x80 to 0xAF. 0x81 has already been assigned within the Digital Television Standard (see Section 5.7.1).

5.6.2.2.2 PMT descriptors

The Ancillary Service specification shall include all pertinent descriptors that are found within the Program Map Table. Specifically, it is recommended that either the `private_data_indicator_descriptor` or the `registration_descriptor`, or both, be included. Although this is not required for a stream with a unique `stream_type` code within this Standard, it will enhance interoperability in the case where the stream is stored outside this Standard, or transmitted in some other network that has its own set of `stream_type` codes.

5.7 Assignment of Identifiers

In this section, those Identifiers and codes that shall have a fixed value are summarized. These include PES Stream IDs and Descriptors. `stream_type` codes from 0x80 to 0xAF shall be reserved for assignment as needed within the Digital Television Standard. `Descriptor_tag` codes from 0x40 to 0xAF shall be reserved for assignment as needed within the Digital Television Standard.

5.7.1 Stream Type

The AC-3 audio `stream_type` shall have the value 0x81.

5.7.2 Descriptors

5.7.2.1 AC-3 audio descriptor

In the digital television system the AC-3 audio descriptor shall be included in the TS_program_map_section. The syntax is given in Table C2 of Annex A of ATSC Standard A/52. There are the following constraints on the AC-3 audio descriptor:

- The value of the descriptor_tag shall be 0x81.
- If textlen exists, it shall have a value of '0x00'.

5.7.2.2 Program smoothing buffer descriptor

The Program Map Table of each program shall contain a smoothing buffer descriptor pertaining to that program in accordance with Section 2.6.30 of ISO/IEC 13818-1. During the continuous existence of a program, the value of the elements of the smoothing buffer descriptor shall not change.

The fields of the smoothing buffer descriptor shall meet the following constraints:

- The field sb_leak_rate shall be allowed to range up to the maximum transport rates specified in Section 7.2.
- The field sb_size shall have a value less than or equal to 2048. The size of the smoothing buffer is thus ≤ 2048 bytes.

5.8 Extensions to the MPEG-2 Systems Specification

This Section covers extensions to the MPEG-2 Systems specification.

5.8.1 Scrambling Control

The scrambling control field within the packet header allows all states to exist in the digital television system as defined in Table C1.

Table C1 Transport Scrambling Control Field

transport_scrambling_control	Function
00	packet payload not scrambled
01	not scrambled, state may be used as a flag for private use defined by the service provider.
10	packet payload scrambled with "even" key
11	packet payload scrambled with "odd" key

Elementary Streams for which the transport_scrambling_control field does not exclusively have the value of '00' for the duration of the program, must carry a CA_descriptor in accordance with Section 2.6.16 of ISO/IEC 13818-1.

The implementation of a digital television delivery system that employs conditional access will require the specification of additional data streams and system constraints.

6. FEATURES OF 13818-1 NOT SUPPORTED BY THIS STANDARD

The transport definition is based on the MPEG-2 Systems standard, ISO/IEC 13818-1; however, it does not implement all parts of the standard. This section describes those elements which are omitted from this Standard.

6.1 Program Streams

This Standard does not include those portions of ISO/IEC 13818-1 and Annex A of ATSC Standard A/52 which pertain exclusively to Program Stream specifications.

6.2 Still Pictures

This Standard does not include those portions of ISO/IEC 13818-1 Transport Stream specification which pertain to the Still Picture model.

7. TRANSPORT ENCODER INTERFACES AND BIT RATES

7.1 Transport Encoder Input Characteristics

The MPEG-2 Systems standard specifies the inputs to the transport system as MPEG-2 elementary streams. It is also possible that systems will be implemented wherein the process of forming PES packets takes place within the video, audio, or other data encoders. In such cases, the inputs to the Transport system would be PES packets. Physical interfaces for these inputs (elementary streams and/or PES packets) may be defined as voluntary industry standards by SMPTE or other standardizing organizations.

7.2 Transport Output Characteristics

Conceptually, the output from the transport system is a continuous MPEG-2 transport stream as defined in this Annex at a constant rate of T_r Mbps when transmitted in an 8 VSB system and $2T_r$ when transmitted in a 16 VSB system where:

$$T_r = 2 \times \frac{188}{208} \frac{312}{313} \frac{684}{286} \times 4.5 = 19.39... \text{ Mbps}$$

and

$$\frac{684}{286} \times 4.5$$

is the symbol rate S_r in Msymbols per second for the transmission subsystem (see Section 4 of Annex D). T_r and S_r shall be locked to each other in frequency.

All transport streams conforming to this Standard shall conform to the ISO/IEC 13818-1 model.

Details of the interface for this output, including its physical characteristics, may be defined as a voluntary industry standard by SMPTE, or other standardizing organizations.

Annex D: RF/Transmission System Characteristics (Normative)

1. SCOPE

This Annex describes the characteristics of the RF/transmission subsystem, which is referred to as the VSB subsystem, of the Digital Television Standard. The VSB subsystem offers two modes: a terrestrial broadcast mode (8 VSB), and a high data rate mode (16 VSB). These are described in separate sections of this document.

2. NORMATIVE REFERENCES

There are no normative references.

3. COMPLIANCE NOTATION

As used in this document, “shall” or “will” 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. TRANSMISSION CHARACTERISTICS FOR TERRESTRIAL BROADCAST

The terrestrial broadcast mode (known as 8 VSB) will support a payload data rate of 19.28... Mbps in a 6 MHz channel. A functional block diagram of a representative 8 VSB terrestrial broadcast transmitter is shown in Figure D1. The input to the transmission subsystem from the transport subsystem is a 19.39... Mbps serial data stream comprised of 188-byte MPEG-compatible data packets (including a sync byte and 187 bytes of data which represent a payload data rate of 19.28... Mbps).

The incoming data is randomized and then processed for *forward error correction* (FEC) in the form of Reed-Solomon (RS) coding (20 RS parity bytes are added to each packet), 1/6 data field interleaving and 2/3 rate trellis coding. The randomization and FEC processes are not applied to the sync byte of the transport packet, which is represented in transmission by a Data Segment Sync signal as described below. Following randomization and forward error correction processing, the data packets are formatted into Data Frames for transmission and Data Segment Sync and Data Field Sync are added.

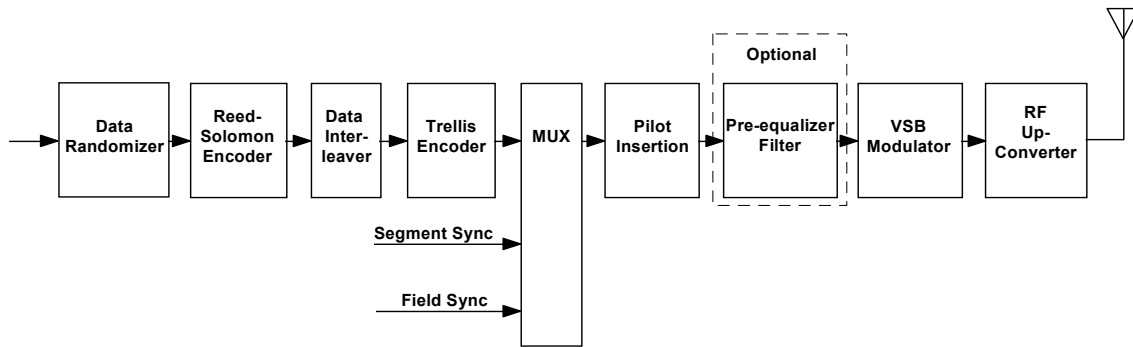


Figure D1 VSB transmitter.

Figure D2 shows how the data are organized for transmission. Each Data Frame consists of two Data Fields, each containing 313 Data Segments. The first Data Segment of each Data Field is a unique synchronizing signal (Data Field Sync) and includes the training sequence used by the equalizer in the receiver. The remaining 312 Data Segments each carry the equivalent of the data from one 188-byte transport packet plus its associated FEC overhead. The actual data in each Data Segment comes from several transport packets because of data interleaving. Each Data Segment consists of 832 symbols. The first 4 symbols are transmitted in binary form and provide segment synchronization. This Data Segment Sync signal also represents the sync byte of the 188-byte MPEG-compatible transport packet. The remaining 828 symbols of each Data Segment carry data equivalent to the remaining 187 bytes of a transport packet and its associated FEC overhead. These 828 symbols are transmitted as 8-level signals and therefore carry three bits per symbol. Thus, $828 \times 3 = 2484$ bits of data are carried in each Data Segment, which exactly matches the requirement to send a protected transport packet:

$$187 \text{ data bytes} + 20 \text{ RS parity bytes} = 207 \text{ bytes}$$

$$207 \text{ bytes} \times 8 \text{ bits/byte} = 1656 \text{ bits}$$

$$2/3 \text{ rate trellis coding requires } 3/2 \times 1656 \text{ bits} = 2484 \text{ bits}$$

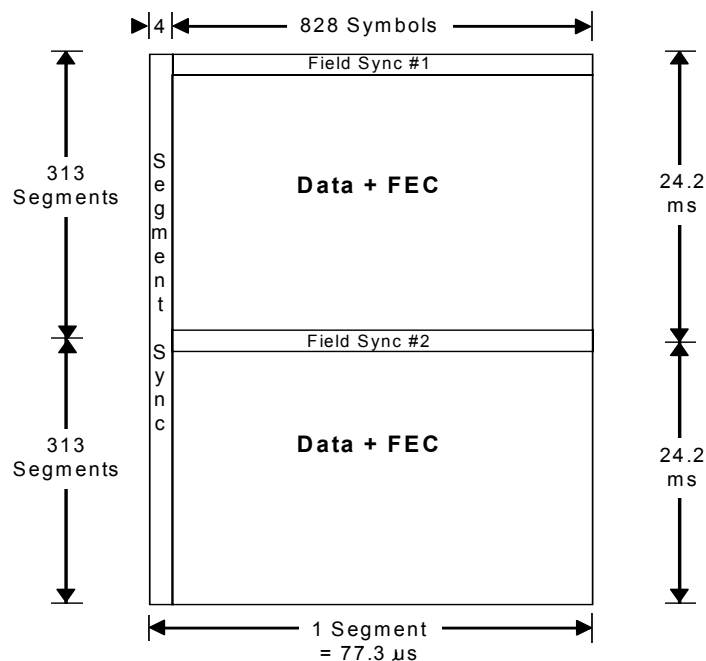


Figure D2 VSB data frame.

The exact symbol rate is given by

$$S_r \text{ (MHz)} = \frac{4.5}{286} \times 684 = 10.76... \text{ MHz}$$

The frequency of a Data Segment is given by

$$f_{seg} = \frac{S_r}{832} = 12.94... \times 10^3 \text{ Data Segments/s}$$

The Data Frame rate is given by

$$f_{frame} = \frac{f_{seg}}{626} = 20.66... \text{ frames/s}$$

The symbol rate S_r and the transport rate T_r (see Section 7.2 of Annex C) shall be locked to each other in frequency.

The 8-level symbols combined with the binary Data Segment Sync and Data Field Sync signals shall be used to suppressed-carrier modulate a single carrier. Before transmission, however, most of the lower sideband shall be removed. The resulting spectrum is flat, except for the band edges where a nominal square root raised cosine response results in 620 kHz transition regions. The nominal VSB transmission spectrum is shown in Figure D3.

At the suppressed-carrier frequency, 310 kHz from the lower band edge, a small pilot shall be added to the signal.

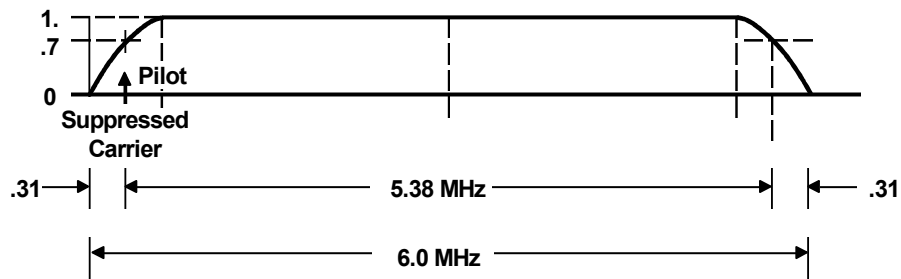


Figure D3 VSB channel occupancy (nominal).

4.1 Channel Error Protection and Synchronization

4.1.1 Prioritization

All payload data shall be carried with the same priority.

4.1.2 Data Randomizer

A data randomizer shall be used on all input data to randomize the data payload (not including Data Field Sync or Data Segment Sync, or RS parity bytes). The data randomizer XORs all the incoming data bytes with a 16-bit maximum length pseudo random binary sequence (PRBS) which is initialized at the beginning of the Data Field. The PRBS is generated in a 16-bit shift register that has 9 feedback taps. Eight of the shift register outputs are selected as the fixed randomizing byte, where each bit from this byte is used to individually XOR the corresponding input data bit. The data bits are XORed MSB to MSB ... LSB to LSB.

The randomizer generator polynomial is as follows:

$$G_{(16)} = X^{16} + X^{13} + X^{12} + X^{11} + X^7 + X^6 + X^3 + X + 1$$

The initialization (pre-load) to F180 hex (load to 1) occurs during the Data Segment Sync interval prior to the first Data Segment.

The randomizer generator polynomial and initialization is shown in Figure D4.

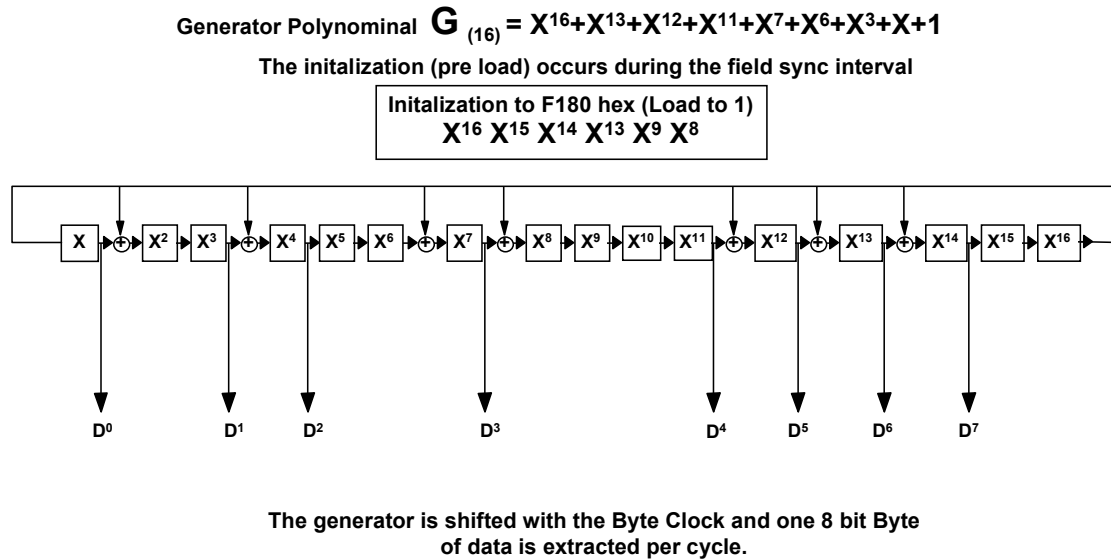


Figure D4 Randomizer polynomial.

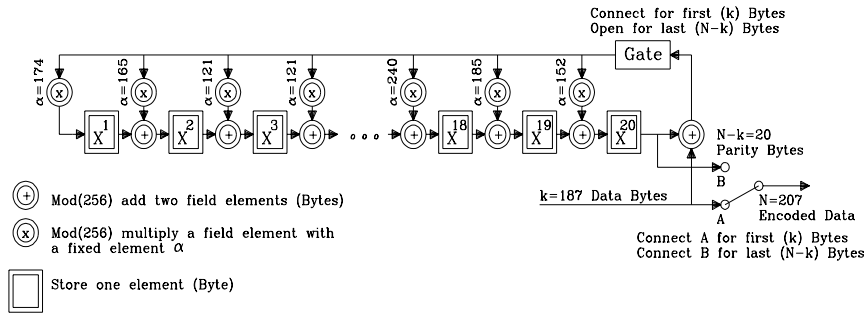
4.1.3 Reed-Solomon Encoder

The RS code used in the VSB transmission subsystem shall be a $t = 10$ (207,187) code. The RS data block size is 187 bytes, with 20 RS parity bytes added for error correction. A total RS block size of 207 bytes is transmitted per Data Segment.

In creating bytes from the serial bit stream, the MSB shall be the first serial bit. The 20 RS parity bytes shall be sent at the end of the Data Segment. The parity generator polynomial and the primitive field generator polynomial are shown in Figure D5.

$$\prod_{i=0}^{i=2t-1} (X + \alpha^i) = X^{20} + X^{19}\alpha^{17} + X^{18}\alpha^{60} + X^{17}\alpha^{99} + X^{16}\alpha^{50} + X^{15}\alpha^{81} + X^{14}\alpha^{163} + X^{13}\alpha^{26} + X^{12}\alpha^{187} + X^{11}\alpha^{202} + X^{10}\alpha^{189} + X^9\alpha^{221} + X^8\alpha^{225} + X^7\alpha^{83} + X^6\alpha^{239} + X^5\alpha^{156} + X^4\alpha^{184} + X^3\alpha^{212} + X^2\alpha^{212} + X^1\alpha^{188} + \alpha^{190}$$

$$= X^{20} + 152X^{19} + 185X^{18} + 240X^{17} + 5X^{16} + 111X^{15} + 99X^{14} + 6X^{13} + 220X^{12} + 112X^{11} + 150X^{10} + 69X^9 + 36X^8 + 187X^7 + 22X^6 + 228X^5 + 198X^4 + 121X^3 + 121X^2 + 165X + 174$$



Primitive Field Generator Polynomial (Galois Field)

$$G(256) = X^8 + X^4 + X^3 + X^2 + 1$$

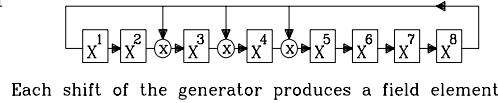


Figure D5 Reed-Solomon (207,187) t=10 parity generator polynomial.

4.1.4 Interleaving

The interleaver employed in the VSB transmission system shall be a 52 data segment (intersegment) convolutional byte interleaver. Interleaving is provided to a depth of about 1/6 of a data field (4 ms deep). Only data bytes shall be interleaved. The interleaver shall be synchronized to the first data byte of the data field. Intra-segment interleaving is also performed for the benefit of the trellis coding process.

The convolutional interleaver is shown in Figure D6.

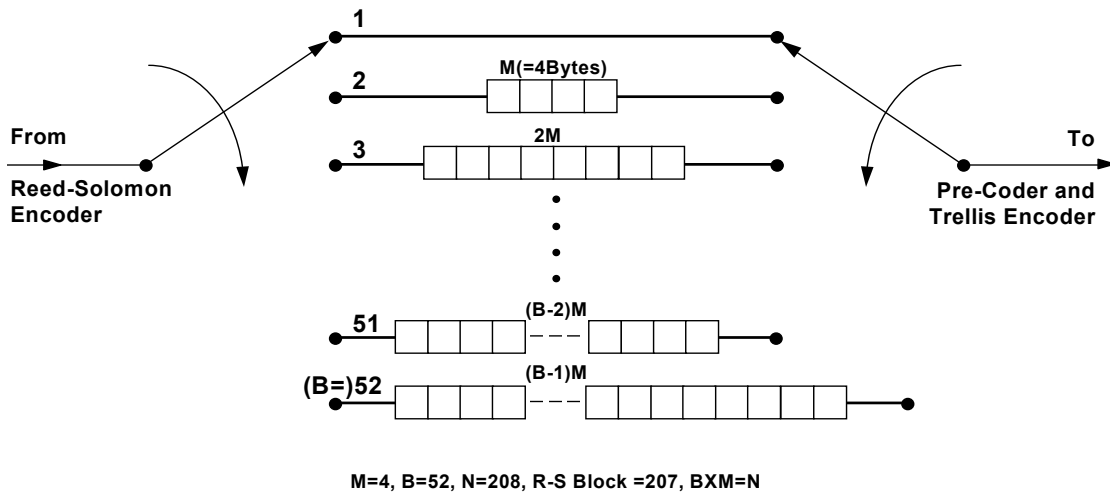


Figure D6 Convolutional interleaver (byte shift register illustration).

4.1.5 Trellis Coding

The 8 VSB transmission sub-system shall employ a $2/3$ rate ($R = 2/3$) trellis code (with one unencoded bit which is precoded). That is, one input bit is encoded into two output bits using a $1/2$ rate convolutional code while the other input bit is precoded. The signaling waveform used with the trellis code is an 8-level (3 bit) one-dimensional constellation. The transmitted signal is referred to as 8 VSB. A 4-state trellis encoder shall be used.

Trellis code intrasegment interleaving shall be used. This uses twelve identical trellis encoders and precoders operating on interleaved data symbols. The code interleaving is accomplished by encoding symbols (0, 12, 24, 36 ...) as one group, symbols (1, 13, 25, 37, ...) as a second group, symbols (2, 14, 26, 38, ...) as a third group, and so on for a total of 12 groups.

In creating serial bits from parallel bytes, the MSB shall be sent out first: (7, 6, 5, 4, 3, 2, 1, 0). The MSB is precoded (7, 5, 3, 1) and the LSB is feedback convolutional encoded (6, 4, 2, 0). Standard 4-state optimal Ungerboeck codes shall be used for the encoding. The trellis code utilizes the 4-state feedback encoder shown in Figure D7. Also shown is the precoder and the symbol mapper. The trellis code and precoder intrasegment interleaver is shown in Figure D8 which feeds the mapper shown in Figure D7. Referring to Figure D8, data bytes are fed from the byte interleaver to the trellis coder and precoder, and they are processed as whole bytes by each of the twelve encoders. Each byte produces four symbols from a single encoder.

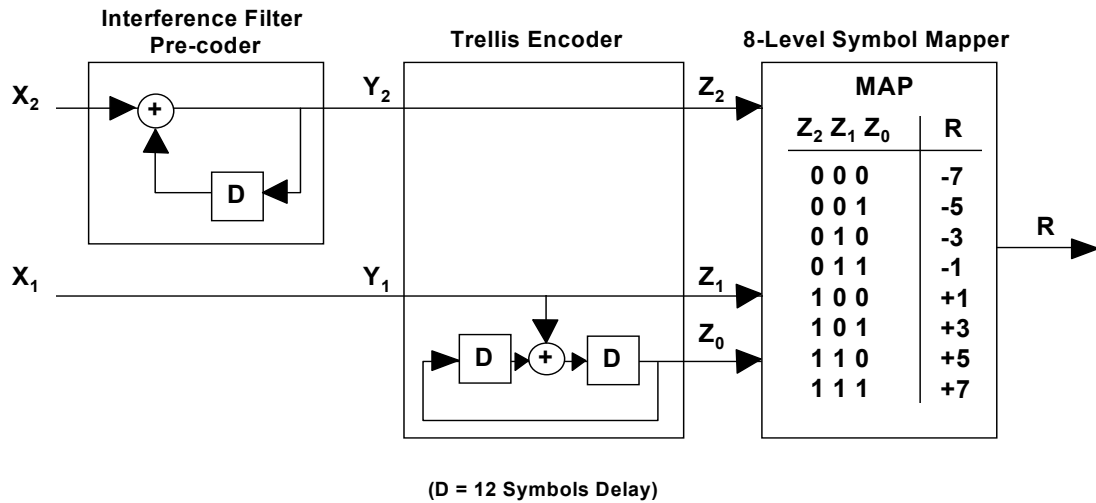


Figure D7 8 VSB trellis encoder, precoder, and symbol mapper.

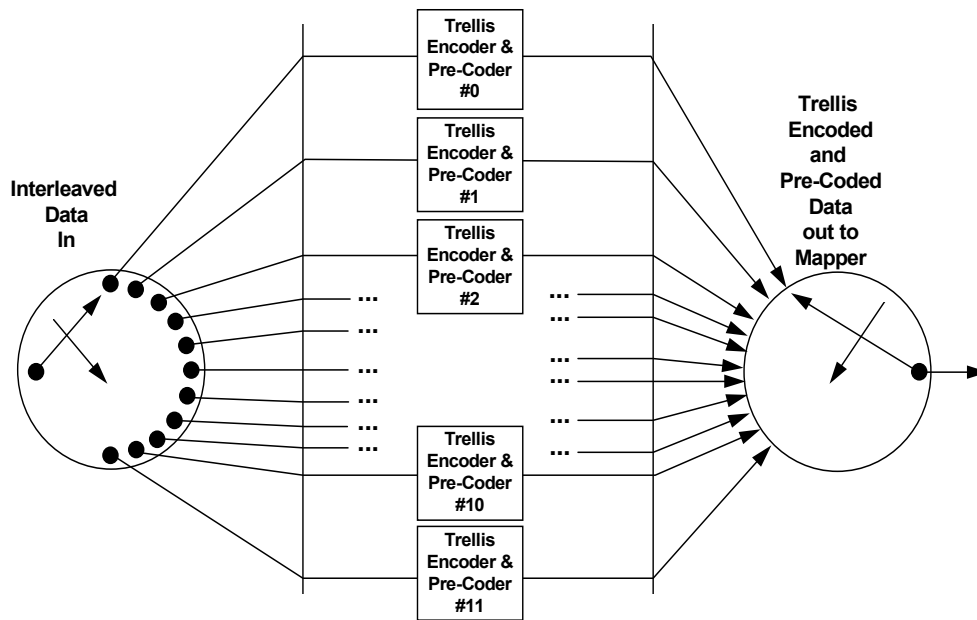


Figure D8 Trellis code interleaver.

The output multiplexer shown in Figure D8 shall advance by four symbols on each segment boundary. However, the state of the trellis encoder shall not be advanced. The data coming out of the multiplexer shall follow normal ordering from encoder 0 through 11 for the first segment of the frame, but on the second segment the order changes and symbols are read from encoders 4 through 11, and then 0 through 3. The third segment reads from encoder 8 through 11 and then 0 through 7. This three-segment pattern shall repeat through the 312 Data Segments of the frame. Table D1 shows the interleaving sequence for the first three Data Segments of the frame.

After the Data Segment Sync is inserted, the ordering of the data symbols is such that symbols from each encoder occur at a spacing of twelve symbols.

Table D1 Interleaving Sequence

Segment	Block 0	Block 1	...	Block 68
0	D0 D1 D2 ... D11	D0 D1 D2 ... D11	...	D0 D1 D2 ... D11
1	D4 D5 D6 ... D3	D4 D5 D6 ... D3	...	D4 D5 D6 ... D3
2	D8 D9 D10 ... D7	D8 D9 D10 ... D7	...	D8 D9 D10 ... D7

A complete conversion of parallel bytes to serial bits needs 828 bytes to produce 6624 bits. Data symbols are created from 2 bits sent in MSB order, so a complete conversion operation yields 3312 data symbols, which corresponds to 4 segments of 828 data symbols. 3312 data symbols divided by 12 trellis encoders gives 276 symbols per trellis encoder. 276 symbols divided by 4 symbols per byte gives 69 bytes per trellis encoder.

The conversion starts with the first segment of the field and proceeds with groups of 4 segments until the end of the field. 312 segments per field divided by 4 gives 78 conversion operations per field.

During segment sync the input to 4 encoders is skipped and the encoders cycle with no input. The input is held until the next multiplex cycle and then fed to the correct encoder.

Table D2 details the byte-to-symbol conversion and the associated multiplexing of the trellis encoders. Segment 0 is the first segment of the field. The pattern repeats every 12 segments; segments 5 through 11 are not shown.

Table D2 Byte to Symbol Conversion, Multiplexing of Trellis Encoders

Symbol	Segment 0			Segment 1			Segment 2			Segment 3			Segment 4		
	Trellis	Byte	Bits	Trellis	Byte	Bits	Trellis	Byte	Bits	Trellis	Byte	Bits	Trellis	Byte	Bits
0	0	0	7,6	4	208	5,4	8	412	3,2	0	616	1,0	4	828	7,6
1	1	1	7,6	5	209	5,4	9	413	3,2	1	617	1,0	5	829	7,6
2	2	2	7,6	6	210	5,4	10	414	3,2	2	618	1,0	6	830	7,6
3	3	3	7,6	7	211	5,4	11	415	3,2	3	619	1,0
4	4	4	7,6	8	212	5,4	0	416	3,2	4	620	1,0
5	5	5	7,6	9	213	5,4	1	417	3,2	5	621	1,0
6	6	6	7,6	10	214	5,4	2	418	3,2	6	622	1,0
7	7	7	7,6	11	215	5,4	3	419	3,2	7	623	1,0
8	8	8	7,6	0	204	5,4	4	408	3,2	8	612	1,0
9	9	9	7,6	1	205	5,4	5	409	3,2	9	613	1,0
10	10	10	7,6	2	206	5,4	6	410	3,2	10	614	1,0
11	11	11	7,6	3	207	5,4	7	411	3,2	11	615	1,0
12	0	0	5,4	4	208	3,2	8	412	1,0	0	624	7,6
13	1	1	5,4	5	209	3,2	9	413	1,0	1	625	7,6
...
19	7	7	5,4	11	215	3,2	3	419	1,0	7	631	7,6
20	8	8	5,4	0	204	3,2	4	408	1,0	8	632	7,6
21	9	9	5,4	1	205	3,2	5	409	1,0	9	633	7,6
22	10	10	5,4	2	206	3,2	6	410	1,0	10	634	7,6
23	11	11	5,4	3	207	3,2	7	411	1,0	11	635	7,6
24	0	0	3,2	4	208	1,0	8	420	7,6	0	624	5,4
25	1	1	3,2	5	209	1,0	9	421	7,6	1	625	5,4
...
31	7	7	3,2	11	215	1,0	3	427	7,6
32	8	8	3,2	0	204	1,0	4	428	7,6
33	9	9	3,2	1	205	1,0	5	429	7,6
34	10	10	3,2	2	206	1,0	6	430	7,6
35	11	11	3,2	3	207	1,0	7	431	7,6
36	0	0	1,0	4	216	7,6	8	420	5,4
37	1	1	1,0	5	217	7,6	9	421	5,4
...
47	11	11	1,0	3	227	7,6
48	0	12	7,6	4	216	5,4
49	1	13	7,6	5	217	5,4
...
95	11	23	1,0
96	0	24	7,6
97	1	25	7,6

Symbol	Segment 0			Segment 1			Segment 2			Segment 3			Segment 4		
	Trellis	Byte	Bits	Trellis	Byte	Bits	Trellis	Byte	Bits	Trellis	Byte	Bits	Trellis	Byte	Bits
...
767	11	191	1,0
768	0	192	7,6
769	1	193	7,6
...
815	11	203	1,0	3	419	7,6	7	623	5,4	11	827	3,2
816	0	204	7,6	4	408	5,4	8	612	3,2	0	816	1,0
817	1	205	7,6	5	409	5,4	9	613	3,2	1	817	1,0
...
827	11	215	7,6	3	419	5,4	7	623	3,2	11	827	1,0

4.1.6 Data Segment Sync

The encoded trellis data shall be passed through a multiplexer that inserts the various synchronization signals (Data Segment Sync and Data Field Sync).

A two-level (binary) 4-symbol Data Segment Sync shall be inserted into the 8-level digital data stream at the beginning of each Data Segment. (The MPEG sync byte shall be replaced by Data Segment Sync.) The Data Segment Sync embedded in random data is illustrated in Figure D9.

A complete segment shall consist of 832 symbols: 4 symbols for Data Segment Sync, and 828 data plus parity symbols. The Data Segment Sync is binary (2-level). The same sync pattern occurs regularly at 77.3 μs intervals, and is the only signal repeating at this rate. Unlike the data, the four symbols for Data Segment Sync are not Reed-Solomon or trellis encoded, nor are they interleaved. The Data Segment Sync pattern shall be a 1001 pattern, as shown in Figure D9.

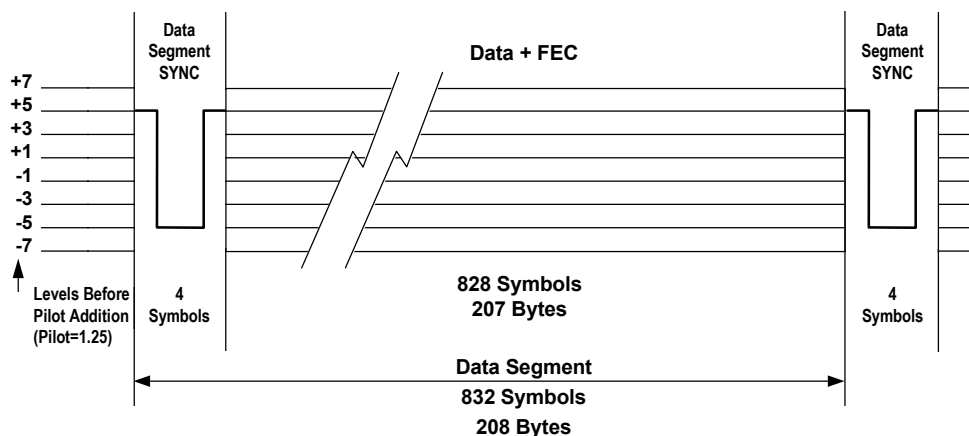
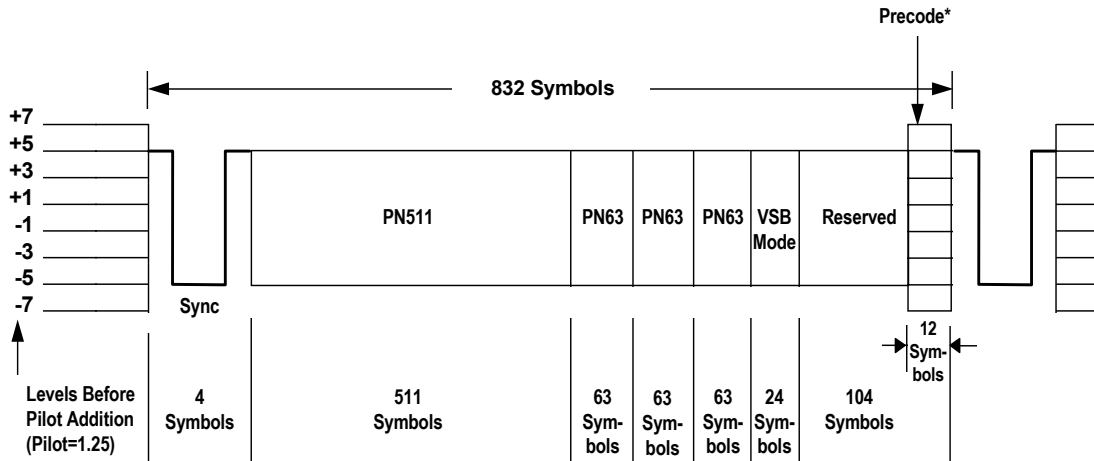


Figure D9 8 VSB data segment.

4.1.7 Data Field Sync

The data are not only divided into Data Segments, but also into Data Fields, each consisting of 313 segments. Each Data Field (24.2 ms) shall start with one complete Data Segment of Data Field Sync, as shown in Figure D10. Each symbol represents one bit of data (2-level). The 832 symbols in this segment are defined below. Refer to Figure D10.



* For trellis coded terrestrial 8 VSB the last 12 symbols of the previous segment are duplicated in the last 12 reserved symbols of the field sync.

Figure D10 VSB data field sync.

4.1.7.1 Sync

This corresponds to Data Segment Sync and is defined as 1001.

4.1.7.2 PN511

This pseudo-random sequence is defined as

$$X^9 + X^7 + X^6 + X^4 + X^3 + X + 1$$

with a pre-load value of 010000000. The sequence is:

```

0000 0001 0111 1111 1100 1010 1010 1110 0110 0110 1000 1000 1001 1110 0001 1101
0111 1101 0011 0101 0011 1011 0011 1010 0100 0101 1000 1111 0010 0001 0100 0111
1100 1111 0101 0001 0100 1100 0011 0001 0000 0100 0011 1111 0000 0101 0100 0000
1100 1111 1110 1110 1010 1001 0110 0110 0011 0111 0111 1011 0100 1010 0100 1110
0111 0001 0111 0100 0011 0100 1111 1011 0001 0101 1011 1100 1101 1010 1110 1101
1001 0110 1101 1100 1001 0010 1110 0011 1001 0111 1010 0011 0101 1000 0100 1101
1111 0001 0010 1011 1100 0110 0101 0000 1000 1100 0001 1110 1111 1101 0110 1010
1100 1001 1001 0001 1101 1100 0010 1101 0000 0110 1100 0000 1001 0000 0001 110
    
```

4.1.7.3 PN63

This pseudo-random sequence is repeated three times. It is defined as

$$X^6 + X + 1$$

with a pre-load value of 100111. The middle PN63 is inverted on every other Data Field Sync. The sequence is:

1110 0100 1011 0111 0110 0110 1010 1111 1100 0001 0000 1100 0101 0011 1101 000

The generators for the PN63 and PN511 sequences are shown in Figure D11.

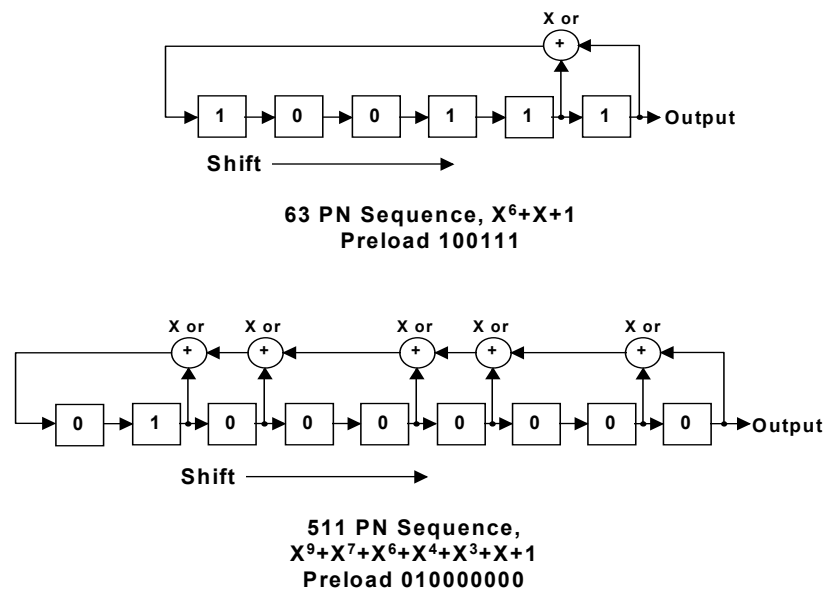


Figure D11 Field sync PN sequence generators.

4.1.7.4 VSB mode

These 24 bits determine the VSB mode for the data in the frame. The first two bytes are reserved. The suggested fill pattern is 0000 1111 0000 1111. The next byte is defined as:

P A B C P A B C

where P is the even parity bit, the MSB of the byte, and A,B, C are the actual mode bits.

P	A	B	C	
0	0	0	0	Reserved
1	0	0	1	Reserved
1	0	1	0	Reserved
0	0	1	1	Reserved
1	1	0	0	16 VSB
0	1	0	1	8 VSB*
0	1	1	0	Reserved
1	1	1	1	Reserved

* In the 8 VSB mode, the preceding bits are defined as :

0 0 0 0 P A B C P A B C 1 1 1 1

4.1.7.5 Reserved

The last 104 bits shall be reserved space. It is suggested that this be filled with a continuation of the PN63 sequence. In the 8 VSB mode, 92 bits are reserved followed by the 12 symbol definition below.

4.1.7.6 Precode

In the 8 VSB mode, the last 12 symbols of the segment shall correspond to the last 12 symbols of the previous segment. All sequences are pre-loaded before the beginning of the Data Field Sync.

Like the Data Segment Sync, the Data Field Sync is not Reed-Solomon or trellis encoded, nor is it interleaved.

4.2 Modulation

4.2.1 Bit-to-Symbol Mapping

Figure D7 shows the mapping of the outputs of the trellis decoder to the nominal signal levels of $(-7, -5, -3, -1, 1, 3, 5, 7)$. As shown in Figure D9, the nominal levels of Data Segment Sync and Data Field Sync are -5 and $+5$. The value of 1.25 is added to all these nominal levels after the bit-to-symbol mapping function for the purpose of creating a small pilot carrier.

4.2.2 Pilot Addition

A small in-phase pilot shall be added to the data signal. The frequency of the pilot shall be the same as the suppressed-carrier frequency as shown in Figure D3. This may be generated in the following manner. A small (digital) dc level (1.25) shall be added to every symbol (data and sync) of the digital baseband data plus sync signal $(\pm 1, \pm 3, \pm 5, \pm 7)$. The power of the pilot shall be 11.3 dB below the average data signal power.

4.2.3 8 VSB Modulation Method

The VSB modulator receives the 10.76 Msymbols/s, 8-level trellis encoded composite data signal (pilot and sync added). The ATV system performance is based on a linear phase raised cosine Nyquist filter response in the concatenated transmitter and receiver, as shown in Figure D12. The system filter response is essentially flat across the entire band, except for the transition regions at each end of the band. Nominally, the roll-off in the transmitter shall have the response of a linear phase root raised cosine filter.

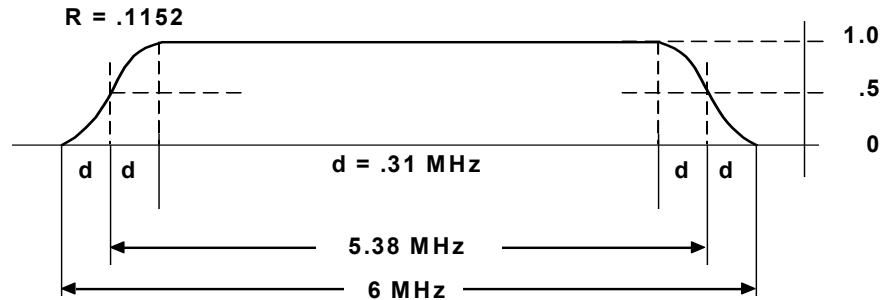


Figure D12 Nominal VSB system channel response (linear phase raised cosine Nyquist filter).

5. TRANSMISSION CHARACTERISTICS FOR HIGH DATA RATE MODE

The high data rate mode trades off transmission robustness (28.3 dB signal-to-noise threshold) for payload data rate (38.57 Mbps). Most parts of the high data rate mode VSB system are identical or similar to the terrestrial system. A pilot, Data Segment Sync, and Data Field Sync are all used to provide robust operation. The pilot in the high data rate mode also is 11.3 dB below the data signal power. The symbol, segment, and field signals and rates are all the same, allowing either receiver to lock up on the other's transmitted signal. Also, the data frame definitions are identical. The primary difference is the number of transmitted levels (8 versus 16) and the use of trellis coding and NTSC interference rejection filtering in the terrestrial system.

The RF spectrum of the high data rate modem transmitter looks identical to the terrestrial system, as illustrated in Figure D3. Figure D13 illustrates a typical data segment, where the number of data levels is seen to be 16 due to the doubled data rate. Each portion of 828 data symbols represents 187 data bytes and 20 Reed-Solomon bytes followed by a second group of 187 data bytes and 20 Reed-Solomon bytes (before convolutional interleaving).

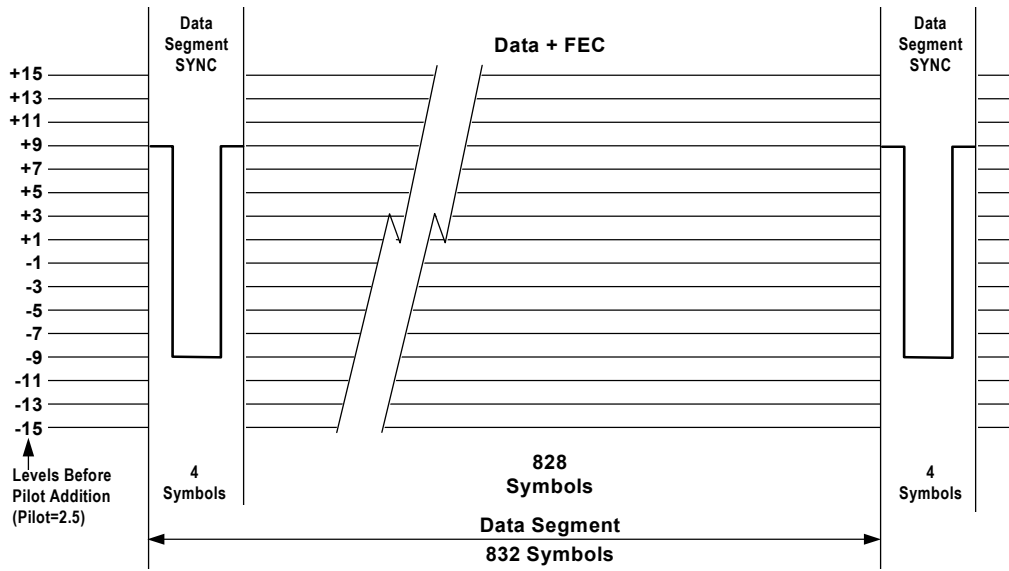


Figure D13 16 VSB data segment.

Figure D14 shows the block diagram of the transmitter. It is identical to the terrestrial VSB system except the trellis coding shall be replaced with a mapper which converts data to multi-level symbols. See Figure D15.

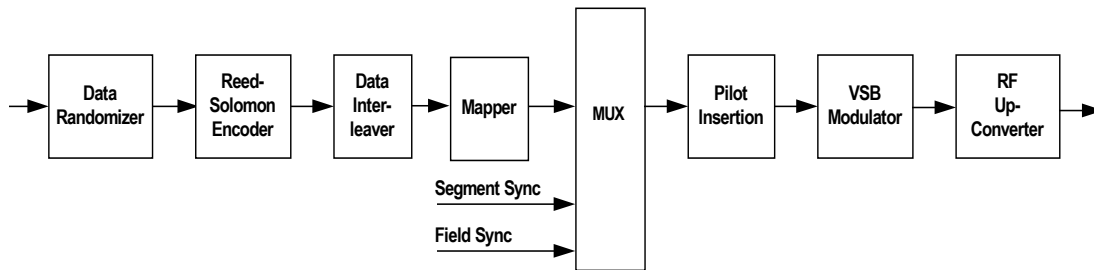


Figure D14 16 VSB transmitter.

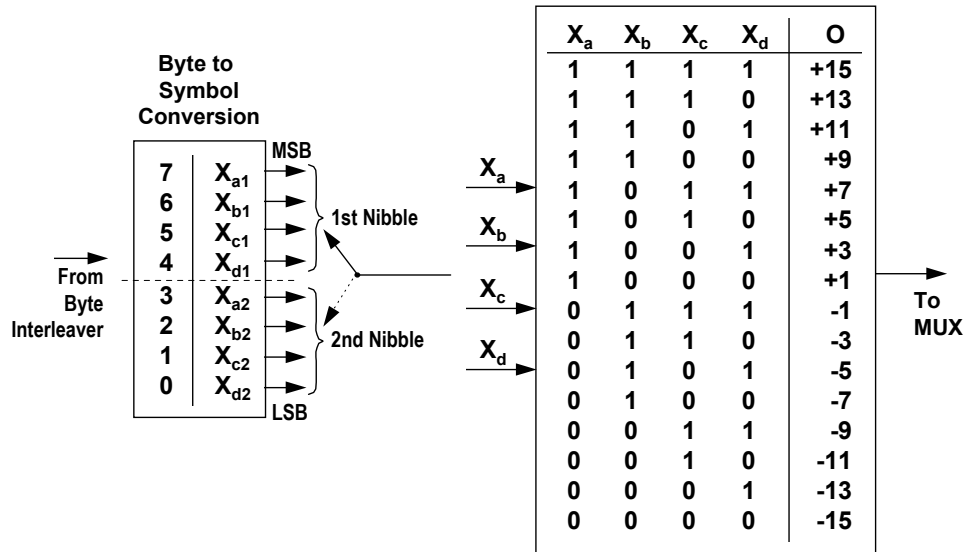


Figure D15 16 VSB mapper.

5.1 Channel Error Protection and Synchronization

5.1.1 Prioritization

See Section 4.1.1.

5.1.2 Data Randomizer

See Section 4.1.2.

5.1.3 Reed-Solomon Encoder

See Section 4.1.3.

5.1.4 Interleaving

The interleaver shall be a 26 data segment inter-segment convolutional byte interleaver. Interleaving is provided to a depth of about 1/12 of a data field (2 ms deep). Only data bytes shall be interleaved.

5.1.5 Data Segment Sync

See Section 4.1.6.

5.1.6 Data Field Sync

See Section 4.1.7.

5.2 Modulation

5.2.1 Bit-to-Symbol Mapping

Figure D15 shows the mapping of the outputs of the interleaver to the nominal signal levels (–15, –13, –11, ..., 11, 13, 15). As shown in Figure D13, the nominal levels of Data Segment Sync and Data Field Sync are –9 and +9. The value of 2.5 is added to all these nominal levels after the bit-to-symbol mapping for the purpose of creating a small pilot carrier.

5.2.2 Pilot Addition

A small in-phase pilot shall be added to the data signal. The frequency of the pilot shall be the same as the suppressed-carrier frequency as shown in Figure D3. This may be generated in the following manner. A small (digital) dc level (2.5) shall be added to every symbol (data and sync) of the digital baseband data plus sync signal (± 1 , ± 3 , ± 5 , ± 7 , ± 9 , ± 11 , ± 13 , ± 15). The power of the pilot shall be 11.3 dB below the average data signal power.

5.2.3 16 VSB Modulation Method

The modulation method shall be identical to that in Section 4, except the number of transmitted levels shall be 16 instead of 8.

Annex E: Receiver Characteristics (Informative)

1. SCOPE

This informative Annex provides material to help readers understand and implement the normative portions of the Digital Television Standard. The normative clauses of the Standard do not specify the design of a receiver. Instead, they specify the transmitted bit stream and RF signal with a thoroughness sufficient to permit the design of a receiver.

Although the normative portions of the Standard are written in the traditional way—by specifying the signal format, not the receiver—the ATSC believes that the introductory phase of this new Standard can be made more orderly by listing some receiver design considerations in this informative Annex. Service providers need assurance that their programs will be correctly processed in all receivers, and receiver manufacturers need assurance that their receivers will function properly with all broadcasts.

This Annex also contains references to existing (both voluntary and mandatory) standards for television receivers and notes work in progress on voluntary industry standards being developed at this time.

2. REFERENCES TO EXISTING OR EMERGING STANDARDS

47 CFR Part 15, FCC Rules.

EIA IS-132, EIA Interim Standard for Channelization of Cable Television.

EIA IS-23, EIA Interim Standard for RF Interface Specification for Television Receiving Devices and Cable Television Systems.

EIA IS-105, EIA Interim Standard for a Decoder Interface Specification for Television Receiving Devices and Cable Television Decoders.

3. COMPLIANCE NOTATION

Compliance with mandatory or voluntary standards and recommended practices for digital television receivers can be inferred only from previous experience with NTSC. Actual standards for digital television receivers have not been developed at this time. As used in this document “appropriate” means that the existing rules for NTSC which are referenced contain most elements of future rules for digital television. Furthermore, the rules may be expanded to cover digital television.

4. STATUS OF RECEIVER STANDARDIZATION ACTIVITIES

4.1 Tuner Performance

The FCC Rules under 47 CFR Part 15 which are applicable to conventional television receivers are expected to be appropriate for digital television receivers.

4.1.1 Noise Figure

The 10 dB noise figure used as a planning factor has been reviewed considering the needs of digital television reception and has been found appropriate.

4.1.2 Channelization Plan for Broadcast and Cable

The cable channelization plan specified in the FCC Rules under 47 CFR Part 15 which are applicable to conventional television receivers are expected to be appropriate for digital television receivers. Broadcast channelization is specified in the FCC Rules under 47 CFR Part 73.

4.1.3 Direct Pickup

The FCC Rules under 47 CFR Part 15 which are applicable to conventional television receivers may be appropriate for digital television receivers as well. Performance characteristics for reception of digital signals, whether standard- or high-definition, have not been developed by the industry. It is expected that direct pickup of a given level will have less effect on digital signals than on NTSC.

4.2 Transport

Significant work for identification of multiple programs within a single digital television channel has not taken place in the industry. It is recommended that a digital television receiver provide appropriate features to assist users in the selection of the desired video program service, if multiple video programs within one channel are offered.

4.3 Decoder Interface

The FCC Rules which are to be adopted for a decoder interface on NTSC receiver advertised as “cable-ready” or “cable-compatible” are expected to be appropriate for digital television receivers. Much work has been done on this interface standard (IS-105) by the Joint Engineering Committee of EIA and NCTA. Although that interface standard is not intended to apply to digital television receivers, it will almost certainly provide a basis for a decoder interface standard applicable to them.

4.4 Digital Data Interface

Work on a digital data interface is being performed by the EIA’s R-4.1 subcommittee on ATV Receiver Interfaces. R-4.1 intends to define a baseband serial digital interface so that devices may exchange packetized data, for example, when a digital VCR is connected to a digital television receiver.

It is recommended that manufacturers of digital television receivers wishing to include a digital data interface give consideration to the interface developed by R4.1.

4.5 Conditional Access Interface

The National Renewable Security System (NRSS) Subcommittee of the Joint Engineering Committee of EIA and NCTA has the responsibility to develop a standard for a plug-in security module. The NRSS standard may be applied in either a standard definition or high definition environment.

It is recommended that manufacturers of digital television receivers wishing to include a conditional access interface give consideration to the NRSS standard developed by the JEC.

4.6 Closed Captioning

Closed captioning for television is covered by the FCC Rules under 47 CFR Part 15 which are presently applied to conventional television receivers. These rules are expected to be appropriate for digital television receivers.

Work on defining the technical standard for closed captioning for the digital television system is being performed by the EIA's R-4.3 subcommittee.

5. RECEIVER FUNCTIONALITY

5.1 Video

It is recommended that a digital television receiver be capable of appropriately decoding and displaying the video scanning formats defined in the Digital Television Standard and described in Table A3 "Compression Format Constraints" in Annex A of this Standard.

5.2 Audio

It is recommended that a digital television receiver be capable of selecting and decoding any audio service described in Section 6 of Annex B of this Standard, subject to the bit rate constraints in Section 5.3 of Annex B of this Standard.

It is recommended that a digital television receiver be capable of normalizing audio levels based on the value of the syntactical element `diaInorm` which is contained in the audio elementary stream.

It is recommended that a digital television receiver be capable of altering reproduced audio levels based on the value of the syntactical element `dynrng` which is contained in the audio elementary stream.

It is recommended that a digital television receiver provide appropriate features to assist users in the selection of program related audio services

Amendment No. 1 to ATSC Digital Television Standard (A/53B)

Advanced Television Systems Committee

1750 K Street, N.W.

Suite 1200

Washington, D.C. 20006

www.atsc.org

Amendment No. 1 to ATSC Digital Television Standard A/53B

This document contains changes to Annex A (Video Systems Characteristics) of the A/53B ATSC Digital Television Standard. These changes are the result of a request by T3 and the Implementation Subcommittee to address issues related to signaling the aspect ratio of “essential picture information” within the MPEG coded frame.

The material presented in Sections 5.2.4.1 and 5.2.4.2 is taken from Annex B of ETSI TR 101 154 V1.4.1, and is reprinted courtesy of ETSI.

1) Annex A, Section 2.1, Normative References

- 1) Add numbers to all references, including “A” designating Annex A.
- 2) Add the SMPTE standards that establish the convention for line numbering for digital pictures as normative references, update publication dates on SMPTE 274M and 296M, move these from informative to normative, and sort the list alphabetically:
 - A10. 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.
 - A11. SMPTE 170M (1999): *Standard for Television—Composite Analog Video Signal, NTSC for Studio Applications*, Society of Motion Picture and Television Engineers, White Plains, N.Y.
 - A12. SMPTE 267M (1995): *Standard for Television—Bit-Parallel Digital Interface, Component Video Signal 4:2:2 16 9 Aspect Ratio*, Society of Motion Picture and Television Engineers, White Plains, N.Y.
 - A13. SMPTE 274M (1998): *Standard for Television—1920 1080 Scanning and Analog and Parallel Digital Interfaces for Multiple Picture Rates*, Society of Motion Picture and Television Engineers, White Plains, N.Y.
 - A14. SMPTE 293M (1996): *Standard for Television—720 483 Active Line at 59.94-Hz Progressive Scan Production, Digital Representation*, Society of Motion Picture and Television Engineers, White Plains, N.Y.
 - A15. SMPTE 296M (2001): *Standard for Television—1280 720 Progressive Image Sample Structure, Analog and Digital Representation and Analog Interface*, Society of Motion Picture and Television Engineers, White Plains, N.Y.

2) Annex A, Section 2.2, Informative References

- 1) Delete SMPTE 274M and 296M from the informative reference list, as they are now normative. Number ITU-R BT.601-4 with “A19” as shown and update it.
- 2) Add a new informative reference to Annex B of DVB ETR-154 to define DVB Active Format Description:

- A16. ETSI TR 101 154 V1.4.1, 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, July 2000.
- 3) Add a new informative reference to Digital TV Group receiver recommendation for aspect ratio signaling:
- A17. ITU-R BT.601-4 (1994): *Encoding Parameters of Digital Television for Studios*.
- A18. *Digital Receiver Implementation Guidelines and Recommended Receiver Reaction to Aspect Ratio Signaling in Digital Video Broadcasting*, Issue 1.2, August 2000, Digital TV Group.
- A19. ITU-R BT.601-5 (1995): *Encoding Parameters of Digital Television for Studios*.
- 4) Add SCTE 21 2001 (Standard for Carriage of NTSC VBI Data in Cable Digital Transport) as an informative reference, as it is referenced in footnote 4.
- A20. SCTE 21 2001 (formerly DVS 053): *Standard for Carriage of NTSC VBI Data in Cable Digital Transport Streams*, Society of Cable Telecommunications Engineers.
-

3) Annex A, Section 5.1.2, Table A3

- 1) In Footnote 2 (footnoted item is “1080” in Table A3), change “last 8 lines” to “bottom 8 lines” to avoid any possible misinterpretation of the last 8 lines of a field (in an interlaced format picture). Footnote 2 to read as shown:

² Note that 1088 lines are actually coded in order to satisfy the MPEG-2 requirement that the coded vertical size be a multiple of 16 (progressive scan) or 32 (interlaced scan). The bottom 8 lines are black, per MPEG rules.

4) Annex A, Section 5.2.2, Picture User Data Syntax

- 1) In the A/53B Standard, only one type of ATSC picture user data is defined, identified with `user_data_type_code` value 0x03. This Amendment adds another type identified with value 0x06. To add this new type, split out the closed caption data (`user_data_type_code` value 0x03) from Table A7. Replace Table A7 with the following table and paragraphs.
- 2) Add an informative note indicating that the bit stream syntax described in Table A6 allows for multiple `user_data(2)` structures within the same video picture. Since the explanation refers to `user_data_type_code` defined in Section 5.2.2, it is added here below the updated Table A7.
- 3) Add a note indicating that receiving devices are expected to process the 32-bit field following `user_data_start_code`. This Amendment adds the possibility that DVB AFD will be present, and it uses a 32-bit value different than the `ATSC_identifier`.

Table A7 Picture User Data Syntax¹

Syntax	No. of Bits	Format
<code>user_data() {</code>		
<code>user_data_start_code</code>	32	bslbf
<code>ATSC_identifier</code>	32	bslbf
<code>user_data_type_code</code>	8	uimsbf
<code>if (user_data_type_code == '0x03')</code>		
<code>cc_data()</code>		
<code>else if (user_data_type_code == '0x06')</code>		
<code>bar_data()</code>		
<code>else {</code>		
<code>while (nextbits() != '0000 0000 0000 0000 0001') {</code>		
<code>ATSC_reserved_user_data</code>	8	
<code>}</code>		
<code>next_start_code()</code>		
<code>}</code>		

In accordance with the bit stream syntax in Table A6, more than one picture user data construct may follow any given picture header. However, no more than one picture user data construct using the same `user_data_type_code` shall follow any given picture header.

Note that picture user data with a 32-bit field following `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 5.2.4). 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 `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) Annex A, Section 5.2.3, Picture User Data Semantics

- 1) Change section title from “Picture User Data Semantics” to “ATSC Picture User Data Semantics.”
- 2) Replace the semantics of `user_data_type_code` with the following paragraphs. Note `ATSC_reserved_user_data` is moved to here (from its location just following `additional_user_data`):

¹ Shaded cells in this table indicate syntactic and semantic additions to the ISO/IEC 13838-2 Standard. Note: `user_data_type_code` values 0x04 and 0x05 are assigned in SCTE 21 2001 [A20].

user_data_type_code An 8-bit value that identifies the type of ATSC user data to follow. Value 0x03 indicates `cc_data()`, value 0x06 indicates `bar_data()`, and other values are either in use in other standards or are reserved for future use.

cc_data() A data structure carrying EIA advanced television closed captioning.

bar_data() A data structure indicating the sizes of letterbox or pillarbox areas within the coded video frame².

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

- 3) Add a level-4 heading just following, with a new table as follows. Changes from the closed caption data in the existing Annex A Table A7 are as indicated.

5.2.3.1 Captioning Data

Table A8 describes the syntax of captioning data.

Table A8 Captioning Data Syntax

Syntax	No. of Bits	Format
<code>cc_data() {</code>		
process_em_data_flag reserved	1	bslbf '1'
process_cc_data_flag	1	bslbf
additional_data_flag	1	bslbf
cc_count	5	uimsbf
em_data reserved	8	bslbf '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		
}		
}		
}		

- 4) There are no plans to transport emergency message data in video user data. Change `process_em_data_flag` and `em_data` in Table A8 to *reserved*, and delete semantic descriptions from the text.

² "Pillarbox" describes a frame that the image fails to fill horizontally, in the same way that a "letterbox" describes a frame that the image fails to fill vertically.

- 5) Add semantics for `additional_cc_data` just following `cc_data_2` as follows:
additional_cc_data Reserved for future ATSC definition.
- 6) Add a level-4 heading, table, and paragraphs just following `additional_cc_data` (keep everything from `cc_data_flag` to `additional_cc_data`):

5.2.3.2 Bar Data

Table A9 describes the syntax of bar data. Bar data 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 frame³.

Bar data shall be carried in the user data of the video Elementary Stream. After each sequence start (and repeat sequence start), the bar data shall be that signalled by the sequence header and sequence display extension parameters. After introduction, bar data shall remain in effect until the next sequence start or until new `bar_data()` is introduced. Receivers should interpret the absence of bar data in a sequence start to mean the rectangular picture area containing useful information extends to the full height and width of the coded frame.

³ In other words, the video is letterboxed (bars above and/or below video) or pillarboxed (bars left and/or right of video).

Table A9 Bar Data Syntax

Syntax	No. of Bits	Format
bar_data() {		
top_bar_flag	1	bslbf
bottom_bar_flag	1	bslbf
left_bar_flag	1	bslbf
right_bar_flag	1	bslbf
reserved	4	'1111'
if (top_bar_flag == 1) {		
marker_bits	2	'11'
line_number_end_of_top_bar	14	uimsbf
}		
if (bottom_bar_flag == 1) {		
marker_bits	2	'11'
line_number_start_of_bottom_bar	14	uimsbf
}		
if (left_bar_flag == 1) {		
marker_bits	2	'11'
pixel_number_end_of_left_bar	14	uimsbf
}		
if (right_bar_flag == 1) {		
marker_bits	2	'11'
pixel_number_start_of_right_bar	14	uimsbf
}		
marker_bits	8	'1111 1111'
while (nextbits() != '0000 0000 0000 0000 0000 0001') {		
additional_bar_data		
}		
}		

Designation of line numbers for line_number_end_of_top_bar and line_number_start_of_bottom_bar is video format-dependent and shall conform to the applicable standard indicated in Table A10.

Table A10 Line Number Designation

Video Format	Applicable Standard
480 Interlaced (analog)	SMPTE 170M [A12]
480 Interlaced 4:3 (digital)	SMPTE 125M [A11]
480 Interlaced 16:9 (digital)	SMPTE 267M [A13]
480 Progressive	SMPTE 293M [A15]
720 Progressive	SMPTE 296M [A16]
1080 Interlaced	SMPTE 274M [A14]
1080 Progressive	SMPTE 274M [A14]

top_bar_flag This flag shall indicate, when set, that the top bar data is present.

bottom_bar_flag This flag shall indicate, when set, that the bottom bar data is present.

left_bar_flag This flag shall indicate, when set, that the left bar data is present.

right_bar_flag This flag shall indicate, when set, that the right bar data is present.

line_number_end_of_top_bar A 14-bit unsigned integer value representing the last line of a horizontal letterbox bar area at the top of the reconstructed frame. Designation of line numbers shall be as defined in Table A10.

line_number_start_of_bottom_bar A 14-bit unsigned integer value representing the first line of a horizontal letterbox bar area at the bottom of the reconstructed frame. Designation of line numbers shall be as defined in Table A10.

pixel_number_end_of_left_bar A 14-bit unsigned integer value representing the last horizontal luminance sample of a vertical pillarbar area at the left side of the reconstructed frame. Pixels shall be numbered from zero, starting with the leftmost pixel.

pixel_number_start_of_right_bar A 14-bit unsigned integer value representing the first horizontal luminance sample of a vertical pillarbar area at the right side of the reconstructed frame. Pixels shall be numbered from zero, starting with the leftmost pixel.

additional_bar_data Reserved for future ATSC definition.

6) Annex A, add new Section 5.2.4

- 1) Add a new section and subsections describing the DVB Active Format Description data and its use.

5.2.4 Active Format Description Data

Active Format Description (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.

The AFD shall be carried 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 A12).

Note: AFD as defined here is identical to that specified in ETSI TR 101 154 V1.4.1 [A16], and is reprinted here with permission.

5.2.4.1 AFD Syntax

Table A11 describes the syntax of the Active Format Description.

Table A11 Active Format Description Syntax

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

5.2.4.2 AFD Semantics

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.

active_format A 4 bit field describing the "area of interest" in terms of its aspect ratio within the coded frame as defined in ISO/IEC 13818-2 [A7].

The active_format is used by the decoder in conjunction with the "source aspect ratio." The source aspect ratio is derived from the "display aspect ratio" (DAR) signaled in the aspect_ratio_information, the horizontal_size, vertical_size, and display_horizontal_size and display_vertical_size if present (see ISO/IEC 13818-2 [A7]):

If sequence_display_extension() is not present, source aspect ratio = DAR

If sequence_display_extension() is present, source aspect ratio =

$$\text{DAR} \times \frac{\text{display_horizontal_size}}{\text{display_vertical_size}} \times \frac{\text{vertical_size}}{\text{horizontal_size}}$$

The combination of source 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 A12 defines the coding of active_format.

Table A12 Active Format

active_format	Aspect Ratio of the “Area of Interest”
'0000' – '0001'	reserved
'0010'	box 16:9 (top)
'0011'	box 14:9 (top)
'0100'	box > 16:9 (center)
'0101' – '0111'	reserved
'1000'	Active format is the same as the coded frame
'1001'	4:3 (center)
'1010'	16:9 (center)
'1011'	14:9 (center)
'1100'	reserved
'1101'	4:3 (with shoot & protect 14:9 center)
'1110'	16:9 (with shoot & protect 14:9 center)
'1111'	16:9 (with shoot & protect 4:3 center)

Illustrations of the various values of active_format may be found in [A18]. Users are strongly encouraged to consult this reference.

5.2.4.3 Recommended Receiver Response to AFD

Receiving device designers are strongly encouraged to study the Recommended Receiver Reaction to Aspect Ratio Signaling [A18] and to design implementations conformant to its recommendations. 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.

7) Annex A, add new informative Section 5.2.5

- 1) Add a new paragraph discussing the relationship between bar data and AFD.

5.2.5 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.