RP 2027-2007

SMPTE RECOMMENDED PRACTICE

AVC Intra-Frame Coding Specification for SSM Card Applications

Table of Contents

Approved
July 10, 2007

Foreword

SMPTE (the Society of Motion Picture and Television Engineers) is an internationally recognized standard developing organization. Headquartered and incorporated in the United States of America, SMPTE has members in over 80 countries on six continents. SMPTE's Engineering Documents, including Standards, Recommended Practices and Engineering Guidelines, are prepared by SMPTE's Technology Committees. Participation in these Committees is open to all with a bona fide interest in their work. SMPTE cooperates closely with other standards-developing organizations, including ISO, IEC and ITU.

SMPTE Engineering Documents are drafted in accordance with the rules given in Part XIII of its Administrative Practices.

SMPTE RP 2027 was prepared by Technology Committee C24.

Introduction

This introduction is informative and does not form an integral part of this Recommended Practice.

This Recommended Practice defines an AVC (Advanced Video Coding) Intra-frame coding specification for SSM (Solid State Media) Card applications using the High 10 Intra profile and High 4:2:2 Intra profile in compliance with ISO/IEC 14496-10 | ITU-T Rec. H.264. A set of basic parameters containing resolution, frame frequency and coded frame size within these profiles is defined along with the stream structure.

1 Scope

This document specifies basic parameters and stream structure for AVC Intra-frame coding using the High 10 Intra profile and the High 4:2:2 Intra profile in compliance with ISO/IEC 14996-10 | ITU-T Rec. H.264. The intended application is acquisition using SSM (Solid State Media) Card defined in SMPTE RP 2006 and subsequent editing processes.

2 Conformance Notation

Normative text is text that describes elements of the design that are indispensable or contains the conformance language keywords: "shall", "should", or "may". Informative text is text that is potentially helpful to the user, but not indispensable, and can be removed, changed, or added editorially without affecting interoperability. Informative text does not contain any conformance keywords.

All text in this document is, by default, normative, except: the Introduction, any section explicitly labeled as "Informative" or individual paragraphs that start with "Note:"

The keywords "shall" and "shall not" indicate requirements strictly to be followed in order to conform to the document and from which no deviation is permitted.

The keywords, "should" and "should not" indicate that, among several possibilities, one is recommended as particularly suitable, without mentioning or excluding others; or that a certain course of action is preferred but not necessarily required; or that (in the negative form) a certain possibility or course of action is deprecated but not prohibited.

The keywords "may" and "need not" indicate courses of action permissible within the limits of the document.

The keyword “reserved” indicates a provision that is not defined at this time, shall not be used, and may be defined in the future. The keyword “forbidden” indicates “reserved” and in addition indicates that the provision will never be defined in the future.

Unless otherwise specified the order of precedence of the types of normative information in this document shall be as follows. Normative prose shall be the authoritative definition. Tables shall be next, followed by formal languages, then figures, and then any other language forms.

3 Normative References

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

ISO/IEC 14496-10, Information Technology — Coding of Audio-Visual Objects — Part 10: Advanced Video Coding | ITU-T Recommendation H.264, Advanced Video Coding for Generic Audio-Visual Services

SMPTE 274M-2005, Television — 1920 x 1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequences for Multiple Picture Rates

SMPTE 296M-2001, Television — 1280 x 720 Progressive Image Sample Structure — Analog and Digital Representation and Analog Interface

SMPTE RP 2006-2006, Solid State Media (SSM) Card Specification

4 Definitions, Symbols and Abbreviations

This document uses same terms defined in the mentioned standards and in other referenced documents.

AVC: ISO/IEC 14496-10 Advanced Video Coding I ITU-T Recommendation H.264

Class: Class xx defines a class of different bit rates depending on the frame rate for the coding specification which provide a similar level of picture quality

coded frame: The VCL NAL units within an AVC access unit containing an AVC coded frame or the VCL NAL units within a pair of consecutive AVC access units containing AVC coded fields of opposite parity

NAL: Network Abstraction Layer

SEI: Supplemental Enhancement Information

SSM: Solid State Media as specified by SMPTE RP 2006

59.94: Term used for 60/1.001

29.97: Term used for 30/1.001

23.98: Term used for 24/1.001

5 Specification of the AVC Intra-Frame Coding

5.1 Overview

This section describes the basic parameters for the AVC Intra-frame coding specification using the High 10 Intra profile and using the High 4:2:2 Intra profile. Class 50 stands for the coding specification using the High 10 Intra profile. Class 100 stands for the coding specification using the High 4:2:2 Intra profile. The set of basic parameters contain the television system, luma samples per line, frame frequency, coded frame size, and some AVC syntax element values within the selected profile and level with the data structure aligned on 512 byte boundaries.

5.2 Class 50 Specification using the High 10 Intra profile

Table 5.1 – Basic parameters

Profile	High 10 Intra	
Television system	1080 line system	720 line system
Luma samples per line after subsampling	1440	960
Level	level 4	level 3.2
Frame frequency	59.94i, 50i, 29.97p, 25p, 23.98p	59.94p, 50p, 29.97p, 25p, 23.98p
Coded frame size	223,232 bytes (29.97 I-frames/s for 59.94i, 29.97p, and 23.98 I-frames/s for 23.98p) 271,360 bytes (25 I-frames/s for 50i, 25p)	111,104 bytes (59.94 I-frames/s for 59.94p, 29.97 I-frames/s for 29.97p, and 23.98 I-frames/s for 23.98p) 135,168 bytes (50 I-frames/s for 50p, and 25 I-frames/s for 25p)
fixed_frame_rate_flag AVC syntax element value	1	1
pic_width_in_mbs_minus1 AVC syntax element value	89	59
frame_mbs_only_flag AVC syntax element value	0 or 1	1
pic_height_in_map_units_minus1 AVC syntax element value	34 * (1 + frame_mbs_only_flag) – 1	44
chroma_format_idc AVC syntax element value	1	1
chroma_loc_info_present_flag AVC syntax element value	0	0

The AVC compression stream in this Class 50 specification using the High 10 Intra profile shall be compliant with the AVC standard using the following profile and levels, High 10 Intra@L4, and High 10 Intra@L3.2. In addition to the requirements of the AVC profile and level, the following restrictions also apply, as summarized in Table 5.1. The basic parameters of the AVC Intra-frame coding specification using the High 10 Intra profile are the following:

- The video sampling structure shall be defined as a ¾ horizontally subsampled luma raster relative to that defined by SMPTE 274M for 1080 line video rasters and SMPTE 296M for 720 line rasters with corresponding 4:2:0 format chroma rasters.

- The supported video sampling structure after resampling shall be 1440x1080 and 960x720. A digital filter example for sampling rate conversion is shown in Annex A.4.

- For 1080 line rasters, frame_crop_top_offset + frame_crop_bottom_offset shall be greater than or equal to 2 * (frame_mbs_only_flag).

- The following HD video signals shall be supported:

  - 1080 line interlaced video at 59.94 Hz field frequency (1080/59.94i) and 50 Hz field frequency (1080/50i).

  - 1080 line progressive video at 29.97 Hz frame frequency (1080/29.97p), 25 Hz frame frequency (1080/25p), and 23.98 Hz frame frequency (1080/23.98p).

  - 720 line progressive video at 59.94 Hz frame frequency (720/59.94p), 50 Hz frame frequency (720/50p), 29.97 Hz frame frequency (1080/29.97p), 25 Hz frame frequency (720/25p), and 23.98 Hz frame frequency (720/23.98p).

- The coded frame size shall be:

  - 223,232 bytes for 1440x1080 (29.97 I-frames/s for 59.94i and 29.97p, and 23.98 I-frames/s for 23.98p).

  - 271,360 bytes for 1440x1080 (25 I-frames/s for 50i and 25p).

  - 111,104 bytes for 960x720 (59.94 I-frames/s for 59.94p, 29.97 I-frames/s for 29.97p, and 23.98 I-frames/s for 23.98p).

  - 135,168 bytes for 960x720 (50 I-frames/s for 50p and 25 I-frames/s for 25p).

5.3 Class 100 Specification using the High 4:2:2 Intra Profile

Table 5.2 – Basic parameters

Profile	High 4:2:2 Intra	
Television system	1080 line system	720 line system
Luma samples per line	1920	1280
Level	level 4.1	
Frame frequency	59.94i, 50i, 29.97p, 25p, 23.98p	59.94p, 50p, 29.97p, 25p, 23.98p
Coded frame size	462,848 bytes (29.97 I-frames/s for 59.94i, 29.97p, and 23.98 I-frames/s for 23.98p) 559,104 bytes (25 I-frames/s for 50i, 25p)	230,912 bytes (59.94 I-frames/s for 59.94p, 29.97 I-frames/s for 29.97p, and 23.98 I-frames/s for 23.98p) 279,040 bytes (50 I-frames/s for 50p, and 25 I-frames/s for 25p)
fixed_frame_rate_flag AVC syntax element value	1	1
pic_width_in_mbs_minus1 AVC syntax element value	119	79
frame_mbs_only_flag AVC syntax element value	0 or 1	1
pic_height_in_map_units_minus1 AVC syntax element value	34 * (1 + frame_mbs_only_flag) – 1	44
chroma_format_idc AVC syntax element value	2	2
chroma_loc_info_present_flag AVC syntax element value	0	0

The AVC compression stream in the Class 100 specification using the High 4:2:2 Intra profile shall be compliant with the AVC standard using the following profile and levels, High 4:2:2 Intra@L4.1. In addition to the requirements of the AVC profile and level, the following restrictions also apply, as summarized in Table 5.2. The basic parameters of the AVC Intra-frame coding specification using the High 4:2:2 Intra profile are the following:

- The video sampling structure shall be defined by SMPTE 274M for 1080 line video rasters and SMPTE 296M for 720 line rasters.
- The supported video sampling structure shall be 1920x1080 and 1280x720.
- For 1080 line rasters, frame_crop_top_offset + frame_crop_bottom_offset shall be greater than or equal to 2 * (frame_mbs_only_flag).

- The following HD video signals shall be supported:
  - 1080 line interlaced video at 59.94 Hz field frequency (1080/59.94i) and 50 Hz field frequency (1080/50i).
  - 1080 line progressive video at 29.97 Hz frame frequency (1080/29.97p), 25 Hz frame frequency (1080/25p), and 23.98 Hz frame frequency (1080/23.98p).
  - 720 line progressive video at 59.94 Hz frame frequency (720/59.94p), 50 Hz frame frequency (720/50p), 29.97 Hz frame frequency (1080/29.97p), 25 Hz frame frequency (720/25p), and 23.98 Hz frame frequency (720/23.98p).
- The coded frame size shall be:
  - 462,848 bytes for 1920x1080 (29.97 I-frames/s for 59.94i and 29.97p, and 23.98 I-frames/s for 23.98p).
  - 559,104 bytes for 1920x1080 (25 I-frames/s for 50i and 25p).
  - 230,912 bytes for 1280x720 (59.94 I-frames/s for 59.94p, 29.97 I-frames/s for 29.97p, and 23.98 I-frames/s for 23.98p).
  - 279,040 bytes for 1280x720 (50 I-frames/s for 50p and 25 I-frames/s for 25p).

NOTE – The higher number for the Class 100 specification refers to using a higher bitrate for the class specification and is independent of the Class 50 specification.

5.4 Ratio between AVC Syntax Element Values for num_units_in_tick and time_scale

As stated in Table 5.1 and Table 5.2, the AVC syntax element value for fixed_frame_rate_flag shall be set to 1. The ratio between the AVC syntax element values for num_units_in_tick and time_scale shall be as shown in Table 5.3 for the various frame rates in the Class 50 and the Class 100 specifications.

Table 5.3 – Ratio between AVC syntax element values for num_units_in_tick and time_scale for Class 50 and Class 100 Specifications

Video Lines/Frequency	time_scale/num_units_in_tick ratio
720/59.94p	120000/1001
720/50p	100/1
1080/59.94i,29.97p 720/29.97p	60000/1001
1080/50i,25p 720/25p	50/1
1080/23.98p 720/23.98p	48000/1001

6 Stream Structure

The bit stream format shall be the byte stream format that is specified in Annex B of the AVC standard. The AVC Annex B byte stream format allows any number of extra zero-valued bytes to precede or follow any NAL unit in the bit stream. Such zero-valued bytes are referred to in the AVC specification as leading_zero_8bits, zero_byte, or trailing_zero_8bits. The presence or absence of these bytes affects only the quantity of data associated with each AVC access unit without affecting the decoding process of the NAL units in the bit stream.

NOTE – Certain constraints on the quantity and alignment of data associated with sequence parameter sets, picture parameter sets, SEI messages, and coded frames are specified below in this section. The insertion of zero-valued bytes by the encoder (as described in the preceding paragraph) is the suggested means of ensuring that the encoded bit stream will obey these constraints.

The following constraints shall be applied to the stream structure and its alignment:

- An access unit delimiter shall be placed at the beginning of every coded frame in a bit stream. The sequence parameter set and the picture parameter set may be placed after the access unit delimiter at the beginning of a bit stream or after the access unit delimiter at the beginning of every coded frame in a bit stream. When the sequence parameter set and the picture parameter set are placed after the access unit delimiter at the beginning of every coded frame in a bit stream, parameters carried in each sequence parameter set and picture parameter set shall be the same in that bit stream. If the sequence parameter set and the picture parameter set are present for a coded frame, the sequence parameter set shall follow the access unit delimiter, the sequence parameter set shall precede the picture parameter set, and the picture parameter set shall precede any optional SEI messages or the space reserved for SEI messages if SEI message NAL units are not present as subsequently defined. If the sequence parameter set and the picture parameter set are not present for a coded frame, no space shall be reserved for them. The length of the access unit delimiter, the sequence parameter set and the picture parameter set shall be a total of 512 bytes.

- Metadata relevant to each picture may be carried in SEI messages. The presence of SEI messages is optional, but if present, shall be multiplexed after the picture parameter set if the sequence parameter set and picture parameter set are present for a coded frame or after the access unit delimiter if the sequence parameter set and picture parameter are not present for a coded frame. Each individual SEI message that may be present shall be a multiple of 512 bytes. If the sequence parameter set and picture parameter set are present for a coded frame, the total length of all SEI messages which may be present shall be a maximum of 512 x 18 bytes for the 1080 line system and 512 x 10 bytes for the 720 line system. If the sequence parameter set and picture parameter set are not present for a coded frame, the total length of the access unit delimiter which shall be present and all SEI messages which may be present shall be a maximum of 512 x 18 bytes for the 1080 line system and 512 x 10 bytes for the 720 line system. When SEI message NAL units are not present, the space for SEI messages shall be reserved by padding.

  NOTE – Therefore, if the sequence parameter set and the picture parameter set are present for a coded frame, there are 512 x 19 bytes before the coded frame for the 1080 line system and 512 x 11 bytes before the coded frame for the 720 line system. If the sequence parameter set and the picture parameter set are not present for a coded frame, there are 512 x 18 bytes before the coded frame for the 1080 line system and 512 x 10 bytes before the coded frame for the 720 line system.

- The total size of each coded frame shall be fixed as specified in section 5.2 or section 5.3.

Annex A (Informative)
Background Information on the Document

This Annex gives background information on the selection of compression parameters specified in this document.

A.1 Intra Picture Only Coding

This specification uses the High 10 Intra profile and the High 4:2:2 Intra profile within the AVC standard. The High 10 Intra profile and High 4:2:2 Intra profile are used in order to fulfill the following requirements for the domain of acquisition by camcorders and subsequent editing processes:

- Ease of frame accurate editing
- Complexity reduction of equipment
- Low coding latency

A.2 Bit Rate Selection

Applications such as acquisition equipment using a SSM Card can support a video file which contains a Digital Interface (DIF) stream defined in SMPTE 314M and SMPTE 370M. Therefore, coded frame sizes for the AVC streams as shown in section 5.2 and section 5.3 are selected so that the total bit rate for an AVC stream fits into a DIF stream as specified in SMPTE 314M and SMPTE 370M. This provides benefits for users and implementers such as a common recording time and a common access performance to a video file.
In a DIF stream, the space allocated for overhead data is greater for the 50Hz system than for the 60Hz system because the overall coded frame size is 20% larger for the 50Hz system. In the AVC stream defined in this Recommended Practice, the space allocated for overhead data is the same for the 50Hz and 60Hz system. Therefore, the coded frame size for the AVC stream in the 50Hz system is increased to more than 20% over the 60Hz system, so that the overall bit rate is approximately the same between the 60Hz system and the 50Hz system.

A.3 Sampling Structure in the Class 50 Specification

Although AVC provides a variety of tools, it could be difficult to achieve adequate picture quality for professional applications by using Intra-only coding to compress the 4:2:2 sampled 1920x1080 video signal or the 4:2:2 sampled 1280x720 video signal to around 50 Mb/s defined in the Class 50 specification. Therefore, the 4:2:0 sampling structure and horizontal subsampling are introduced in the Class 50 specification to reduce artifacts caused by compression. An example of a digital filter using for the sampling rate conversion used in the Class 50 specification is shown in section A.4. The class 100 specification allows for 4:2:2 and full horizontal sampling. No subsampling is performed on the 4:2:2 sampled 1920x1080 video signal or the 4:2:2 sampled 1280x720 video signal within the Class 100 specification.

A.4 Digital Filter Example for Sampling Rate Conversion

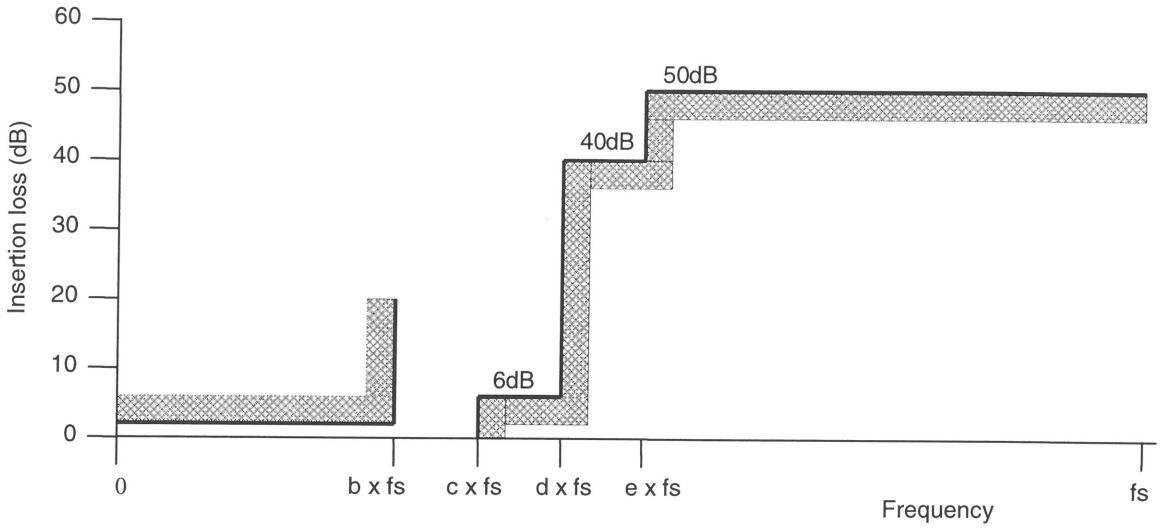

Figure A.1 – Template for insertion loss frequency characteristic

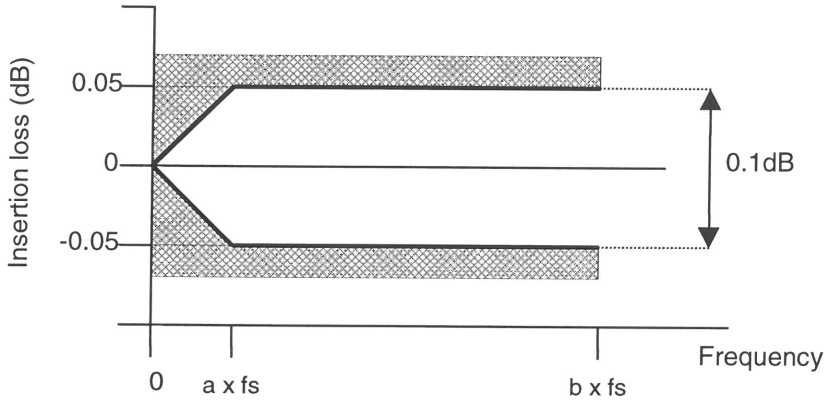

Figure A.2 – Pass band ripple tolerance

Table A.1 – Parameters for Digital Filter

Video Lines/Frequency	fs	Component	a	b	c	d	e
1080/59.94i,29.97p,23.98p 720/50p,25p	74.25/1.001 MHz	Y	0.05	0.25	0.375	0.50	0.60
		C_B, C_R	0.025	0.125	0.1875	0.25	0.30
1080/50i,25p 720/50p,25p	74.25 MHz	Y	0.05	0.25	0.375	0.50	0.60
		C_B, C_R	0.025	0.125	0.1875	0.25	0.30

Annex B (Informative)
Encoder Implementation Guidelines

This Annex gives suggestions for choices of encoding tools using this Recommended Practice. The information in this annex is strictly for purposes of providing advice on encoding practices, and in particular is not intended to specify a relaxation of decoder conformance requirements.

In actual encoder implementation, supporting all of the AVC coding tools involves substantial complexity. Therefore, acquisition equipment, which requires the reduction of implementation complexity, can use encoding tools as suggested by the parameter values shown in Table B.1 for the Class 50 specification and Table B.2 for the Class 100 specification. For the Class 50 specification, the use of Context Adaptive Binary Arithmetic Coding (CABAC) by selecting entropy_coding_mode_flag = 1 is suggested in order to provide for maximization of picture quality within the specified bit rate constraint. For the Class 100 specification, the use of Context Adaptive Variable Length Coding (CAVLC) by selecting entropy_coding_mode_flag = 0 is suggested in order to improve decoding speed for real-time (e.g., software-based) decoding and in consideration that the quality improvement offered by the use of CABAC could be less critical when considering the higher picture quality that is already attainable by using the higher bit rate that applies in Class 100 systems (relative to the bit rate used in Class 50 systems). In addition, specific applications where higher complexity, such as high performance processing, justifies using CABAC to achieve higher picture quality, CABAC might be selected for the Class 100 specification.

Table B.1 – Parameter Value Recommendations for Class 50 Specification

Encoding Parameter	1440x1080, 4:2:0		960x720, 4:2:0
	Interlaced	Progressive	
frame_mbs_only_flag	0	1	1
mb_adaptive_frame_field_flag	1	N.A.	N.A.
entropy_coding_mode_flag	1 (CABAC)		
transform_8x8_mode_flag	1		
frame_cropping_flag	1		0
frame_crop_left_offset	0		N.A.
frame_crop_right_offset	0		N.A.
frame_crop_top_offset	0		N.A.
frame_crop_bottom_offset	2	4	N.A.

Table B.2 – Parameter Value Recommendations for Class 100 Specification

Encoding Parameter	1920x1080, 4:2:2		1280x720, 4:2:2
	Interlaced	Progressive	
frame_mbs_only_flag	0	1	1
mb_adaptive_frame_field_flag	1	N.A.	N.A.
transform_8x8_mode_flag	1		
frame_cropping_flag	1		0
frame_crop_left_offset	0		N.A.
frame_crop_right_offset	0		N.A.
frame_crop_top_offset	0		N.A.
frame_crop_bottom_offset	4	8	N.A.

NOTE – N.A. in Table B.1 and Table B.2 stands for Not Applicable. The syntax elements having entries marked as "N.A." are not present in the AVC bit stream when other related syntax elements have the value shown in these tables. These syntax elements typically have inferred values that are specified in the AVC specification in such cases. The use of "N.A." indicates that the value specified for such inference in the AVC specification applies.

Annex C (Informative)
Bibliography

SMPTE 314M-2005, Television — Data Structure for DV-Based Audio, Data and Compressed Video — 25 and 50 Mb/s

SMPTE 370M-2006, Television — Data Structure for DV-Based Audio, Data and Compressed Video at 100Mb/s — 1080/60i, 1080/50i, 720/60p, 720/50p