IEC TS 62871-2

Edition 1.0 2019-12

TECHNICAL SPECIFICATION

colour inside

Professional video storage products – Tape-less camera recorder using MXF file format – Encoding guidelines – Part 2: Mapping MPEG-2 and AVC Streams into MXF

IEC TS 62871-2:2019-12(en)

IEC TS 62871-2

Edition 1.0 2019-12

TECHNICAL SPECIFICATION

colour inside

Professional video storage products – Tape-less camera recorder using MXF file format – Encoding guidelines – Part 2: Mapping MPEG-2 and AVC Streams into MXF

INTERNATIONAL
ELECTROTECHNICAL
COMMISSION

ICS 33.160; 35.040.40

ISBN 978-2-8322-7711-9

Warning! Make sure that you obtained this publication from an authorized distributor.

CONTENTS

INTERNATIONAL ELECTROTECHNICAL COMMISSION

____

PROFESSIONAL VIDEO STORAGE PRODUCTS – TAPE-LESS CAMERA RECORDER USING MXF FILE FORMAT – ENCODING GUIDELINES –

Part 2: Mapping MPEG-2 and AVC Streams into MXF

FOREWORD

1) The International Electrotechnical Commission (IEC) is a worldwide organization for standardization comprising all national electrotechnical committees (IEC National Committees). The object of IEC is to promote international co-operation on all questions concerning standardization in the electrical and electronic fields. To this end and in addition to other activities, IEC publishes International Standards, Technical Specifications, Technical Reports, Publicly Available Specifications (PAS) and Guides (hereafter referred to as "IEC Publication(s)"). Their preparation is entrusted to technical committees; any IEC National Committee interested in the subject dealt with may participate in this preparatory work. International, governmental and non-governmental organizations liaising with the IEC also participate in this preparation. IEC collaborates closely with the International Organization for Standardization (ISO) in accordance with conditions determined by agreement between the two organizations.
2) The formal decisions or agreements of IEC on technical matters express, as nearly as possible, an international consensus of opinion on the relevant subjects since each technical committee has representation from all interested IEC National Committees.
3) IEC Publications have the form of recommendations for international use and are accepted by IEC National Committees in that sense. While all reasonable efforts are made to ensure that the technical content of IEC Publications is accurate, IEC cannot be held responsible for the way in which they are used or for any misinterpretation by any end user.
4) In order to promote international uniformity, IEC National Committees undertake to apply IEC Publications transparently to the maximum extent possible in their national and regional publications. Any divergence between any IEC Publication and the corresponding national or regional publication shall be clearly indicated in the latter.
5) IEC itself does not provide any attestation of conformity. Independent certification bodies provide conformity assessment services and, in some areas, access to IEC marks of conformity. IEC is not responsible for any services carried out by independent certification bodies.
6) All users should ensure that they have the latest edition of this publication.
7) No liability shall attach to IEC or its directors, employees, servants or agents including individual experts and members of its technical committees and IEC National Committees for any personal injury, property damage or other damage of any nature whatsoever, whether direct or indirect, or for costs (including legal fees) and expenses arising out of the publication, use of, or reliance upon, this IEC Publication or any other IEC Publications.
8) Attention is drawn to the Normative references cited in this publication. Use of the referenced publications is indispensable for the correct application of this publication.
9) Attention is drawn to the possibility that some of the elements of this IEC Publication may be the subject of patent rights. IEC shall not be held responsible for identifying any or all such patent rights.

The main task of IEC technical committees is to prepare International Standards. In exceptional circumstances, a technical committee may propose the publication of a technical specification when

- the required support cannot be obtained for the publication of an International Standard, despite repeated efforts, or
- the subject is still under technical development or where, for any other reason, there is the future but no immediate possibility of an agreement on an International Standard.

Technical specifications are subject to review within three years of publication to decide whether they can be transformed into International Standards.

IEC TS 62871-2, which is a technical specification, has been prepared by technical area 6: Storage media, storage data structure, storage systems and equipment, of IEC technical committee 100: Audio, video and multimedia systems and equipment.

The text of this technical specification is based on the following documents:

Enquiry draft	Report on voting
100/3152/DTS	100/3218/RVDTS

Full information on the voting for the approval of this technical specification can be found in the report on voting indicated in the above table.

This document has been drafted in accordance with the ISO/IEC Directives, Part 2.

A list of all parts in the IEC 62871 series, published under the general title *Professional video storage products – Tape-less camera recorder using MXF file format – Encoding guidelines*, can be found on the IEC website.

The committee has decided that the contents of this document will remain unchanged until the stability date indicated on the IEC website under "http://webstore.iec.ch" in the data related to the specific document. At this date, the document will be

- reconfirmed,
- withdrawn,
- replaced by a revised edition, or
- amended.

A bilingual version of this publication may be issued at a later date.

IMPORTANT – The 'colour inside' logo on the cover page of this publication indicates that it contains colours which are considered to be useful for the correct understanding of its contents. Users should therefore print this document using a colour printer.

INTRODUCTION

The professional camera recorder has evolved from a traditional tape-based system into a file-based system, taking advantage of recent advances in information technology. Instead of using conventional magnetic tape as the recording medium, video and audio streams can now be stored as files that can be read directly by a personal computer (PC).

Several file format specifications exist, and most broadcasters are using the Material eXchange Format (MXF), which has been standardized by the Society of Motion Picture and Television Engineers (SMPTE). As reported in IEC TR 62712:2011, the MXF file format has been adopted for various types of professional tape-less camera recorders. MXF is being used by many broadcast stations around the world. Since the MXF file format provides a multiplicity of functions and options in order to satisfy the needs of various applications in a range of situations, it is important to address interoperability issues between equipment. Therefore, it is essential for interoperability that there is an appropriate Technical Specification that specifies guidelines for MXF implementations and operational usage.

The IEC 62871 series gives encoding guidelines for professional tape-less camera recorders using the MXF file format to ensure interoperability.

The IEC 62871 series currently consists of:

– IEC 62871-1, which gives guidelines for MXF operational patterns for professional tape-less camera recorders and also outlines the general parts of the MXF file format.
– IEC 62871-2, which gives guidelines for mapping MPEG-2 and AVC Streams into MXF files which are used for professional tape-less camera recorders.

PROFESSIONAL VIDEO STORAGE PRODUCTS – TAPE-LESS CAMERA RECORDER USING MXF FILE FORMAT – ENCODING GUIDELINES –

Part 2: Mapping MPEG-2 and AVC Streams into MXF

1 Scope

This part of IEC 62871, which is a technical specification, specifies implementation guidelines for mapping MPEG-2 video and AVC streams into MXF file format for professional tape-less camera recorders.

The guidelines are applicable to the creation of an MXF file in professional tape-less camera recorders. They are also applicable for content management software and to equipment that supports MXF files generated by professional tape-less camera recorders.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 13818-2:2013, *Information technology – Generic coding of moving pictures and associated audio information – Part 2: Video*

ISO/IEC 14496-10:2014, *Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding*

SMPTE ST 377-1:2011, *Material Exchange Format (MXF) – File Format Specification*

SMPTE ST 381-2:2018, *Material Exchange Format (MXF) – Mapping MPEG Streams into the MXF Constrained Generic Container*

SMPTE ST 381-3:2017, *Material Exchange Format (MXF) – Mapping AVC Streams into the MXF Constrained Generic Container*

3 Terms and definitions

For the purposes of this document, the terms, definitions and abbreviated terms given in ISO/IEC 13818-2, ISO/IEC 14496-10, SMPTE ST 377-1, SMPTE ST 381-2 and SMPTE ST 381-3 apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- IEC Electropedia: available at http://www.electropedia.org/
- ISO Online browsing platform: available at http://www.iso.org/obp

4 Mapping of MPEG-2 video and AVC streams into MXF

4.1 Video codec

The MXF file body, which contains video, sound and data is specified with various types of video codec. As reported in IEC TR 62712, two types of video codec, MPEG-2 defined in ISO/IEC 13818-2 and MPEG-4 AVC/H.264 defined in ISO/IEC 144960-10, are widely used in professional tape-less camera recorders.

ISO/IEC 13812-2, the MPEG-2 video specification, enables the compression of picture-based video. In this specification, for progressive video, a picture is identical to a frame, while for interlaced video a picture can refer to a frame or to the top field or the bottom field of the frame depending on the context.

ISO/IEC 14496-10, the MPEG-4 AVC/H.264 specification, enables more efficient compression than MPEG-2 video compression generally.

In this guideline, the term "picture" is used as a collective term for a frame or a field, as defined in ISO/IEC 14496-10.

In both specifications, progressive video is encoded frame-based, and interlaced video is encoded field-based. A sequence of macroblocks is defined as a "slice". A picture can be divided into multiple slices, or a picture can be constituted by one slice. Further, the slice can be divided into multiple blocks.

4.2 Video codec mapping

The wrapping specification of an MPEG-2 video stream is specified in SMPTE ST 381-2, while for an AVC stream it is specified in SMPTE ST 381-3. A sequence of pictures is KLV-coded as defined in SMPTE ST 336. The MPEG-2 video and AVC streams are KLV-wrapped using the MPEG Picture Element Key defined in SMPTE ST 381-2. One type of KLV wrapping, defined as "frame wrapping", is specified for the MPEG body structure and is described in this subclause.

As described in SMPTE ST 379-2, frame wrapping has one or more content packages in the essence container. An example of frame wrapping of MPEG-2 video using MXF Operational Pattern 1a (OP-1a) is shown in Figure 1. Each Content Package has the duration of one MPEG video access unit (AU). Specifying the duration through an MPEG video AU determines where the MPEG headers will be found. Through these MPEG headers, the picture type can be determined as shown in Annex C of SMPTE ST 381-2:2018.

The frame wrapping method is intended to enable retrieval of individual AUs and the corresponding field/frame within them by MXF applications which process at the KLV level. This can be particularly useful for applications that support frame-based access in order to edit or play back randomly.

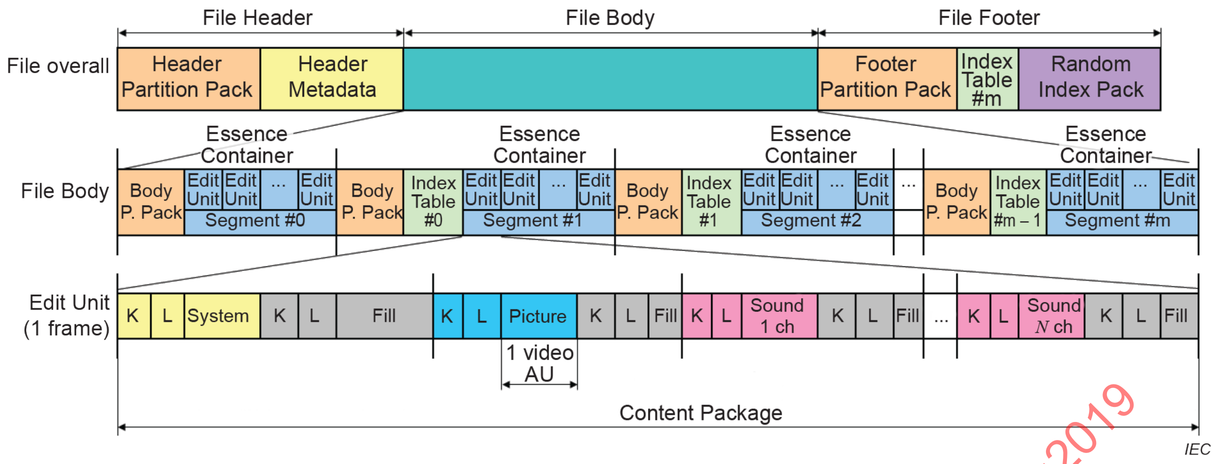

Figure 1 – Example of MPEG-2 video into MXF OP-1a File Body

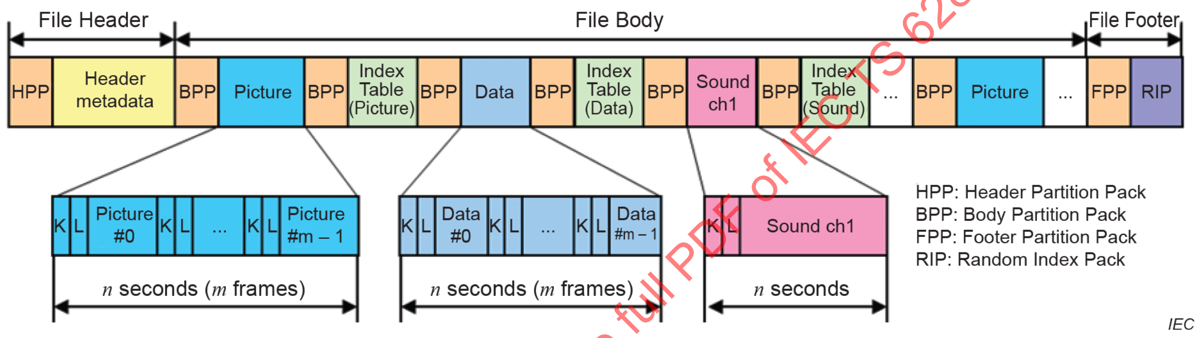

Figure 2 – Example of AVC Stream into MXF OP-1b File Body

An implementation example of mapping an AVC stream into MXF using Operational Pattern 1b (OP-1b) is shown in Figure 2. IEC TS 62871-1 describes Operational Pattern Atom (OP-Atom), which maps independent essences into separate files, but in Figure 2, OP-1b is used in order to map multiple essences independently into a single file.

The MPEG specification provides signalling for the number of pixels in the picture, the field or frame rate and other parameters, including the aspect ratio of the pixels. Many of these parameters will be copied into the MXF Essence Descriptor.

5 Recommended constraints for professional tape-less camera recorders

5.1 General

This guideline shall comply with SMPTE ST381-2 and SMPTE ST381-3, and describes recommended implementation constraints in Subclauses 5.2 to 5.8.

5.2 Frame wrapping

As described in 4.1, the MPEG-2 video and AVC specifications enable the compression of a picture-based video stream. SMPTE ST381-2 and ST 381-3 provide frame wrapping as one of the possible methods of wrapping MPEG-2 video and AVC streams into the MXF file body. In this guideline, frame wrapping should be used in all cases, even if the picture is field-based interlaced video. In the case of Long GOP compression, video compression is performed on a field-by-field basis from the viewpoint of compression efficiency, and it is recommended that the frame-wrapped KLV packet should contain a pair of coded fields.

5.3 Edit Unit size

SMPTE ST 377-1 specifies the Edit Unit which defines the duration of contents components. For the purpose of video editing, the Edit Unit is recommended to be 1 frame, even if Long GOP is used or the interlaced video is field-encoded. In this way, video editing with the resulting MXF file can be performed on a one-frame basis.

5.4 Segment size of Index table

SMPTE ST 377-1 does not specify a limitation on the size of the "Index Entry Array" belonging to the Index Table, although SMPTE 377M-2004 specified the maximum size as 65 535 bytes. Therefore, it is recommended to segment the Index Table in case the Index Entry Array exceeds 65 535 bytes in order to maintain interoperability with SMPTE 377M-2004-compliant MXF decoders.

Furthermore, when segmenting the Index Table into multiple Index Table Segments, since there is no provision to specify the number of Edit Units managed by a particular Index Table Segment, it could be recorded with a different number of Edit Units to those of other Index Table Segments, as shown in Figure 3. In such a case, when the playback device needs to access a certain Edit Unit in order to trick-play (for example, fast-forward playback) from anywhere in the stream, the location of the desired Edit Unit has to be obtained by parsing sequentially from the first Index Table Segment, thereby lowering processing efficiency and taking longer to perform.

Therefore, it is recommended to record each Index Table Segment with the same size, as shown in Figure 4. This can reduce the processing time because it is necessary to obtain only the size of the first Index Table Segment in order to determine the offset of the Index Table Segment where the desired Edit Unit exists.

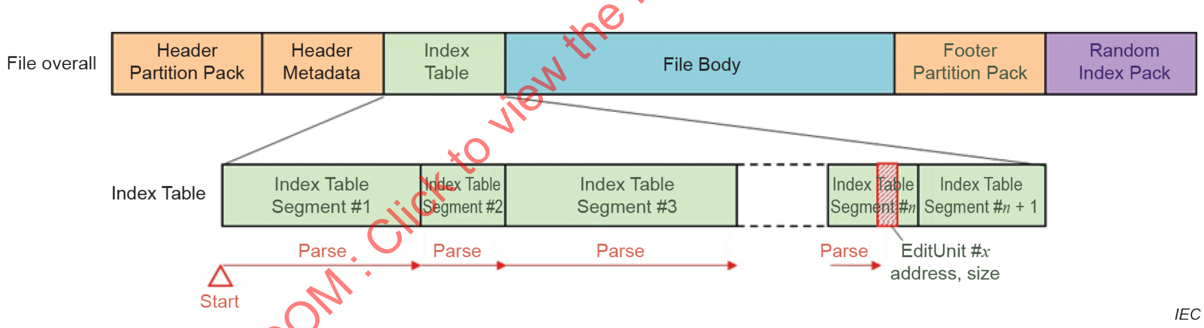

Figure 3 – Example of Index Table Segment with different size

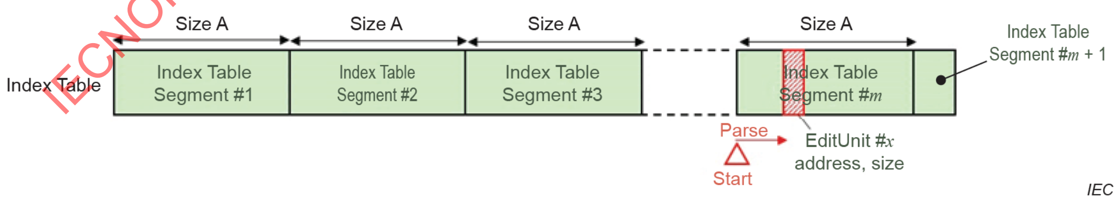

Figure 4 – Example of Index Table Segment with same size

5.5 Codec constraints

This subclause describes a recommended structure of MPEG-2 and AVC Long GOP streams for MXF wrapping to provide fast access to target pictures and to reduce decoder complexity. In this subclause, codec constraints on MPEG-2 and AVC Long GOP streams are recommended by referring to SMPTE ST 381-3:2017, Annex C.

- All P Pictures should consist of only P-Slices and all B Pictures or Br Pictures should consist of only B-Slices, i.e. a picture should not consist of different slice types. Also, a field should not consist of different slice types when field coding is used for interlaced video.
- The first access unit in coded order should be an IDR or non-IDR I Picture in GOP.
- For each picture in the GOP, the corresponding Key Picture should be present within the same GOP or the immediately preceding GOP.
- A P Picture should not refer to a B Picture.
- The coded order for all I and/or P Pictures should be the same as their display order.
- The coded order for non-reference B Pictures that are contiguous in the bit stream should be the same as their display order.

The technology of variable speed decoding and editing of a stream developed for MPEG-2 can be also used for an AVC/H.264 video stream because these constraints make the concept of GOP equivalent to that in MPEG-2 except for the existence of Br Pictures.

5.6 Placement of parameter set (SPS/PPS)

Sequence parameter set (hereinafter referred to as SPS) and picture parameter set (hereinafter referred to as PPS) are parameters that include information necessary for decoding the AVC compressed stream. The implementation scheme of SPS and PPS in the AVC stream affects the performance of playback and editing of the AVC stream in the MXF file recorded by the camera recorder.

The implementation provision of SPS and PPS in the AVC standard is very flexible and only requires being capable of acquiring SPS and PPS at the time of decoding the picture. SMPTE ST 381-3 defines metadata elements in the AVC Sub Descriptor in the MXF Header Metadata to identify the implementation scheme of SPS and PPS in the AVC stream included in the MXF file. By referring to these metadata elements, the decoder can identify the implementation scheme of SPS and PPS within the AVC stream. SMPTE ST 381-3 does not specify any recommended implementations or constraints of SPS and PPS except for multiplexing SPS and PPS in the AVC stream.

This subsection describes the guidelines for SPS and PPS implementations in tape-less camera recorders using AVC compression and gives features according to each implementation example.

NOTE 1 Although PPS is referred to by each slice unit, as described in 5.5, mixing different types of slices within a picture is not recommended in this guideline. In the following implementation examples, for simplicity of explanation, it is assumed that the same PPS is referred to in units of pictures.

NOTE 2 In the drawings of the implementation examples, for simplicity of the drawings, elements other than SPS, PPS and compressed picture are omitted as components of the AVC stream.

- SPS and PPS necessary for decoding each picture are placed at the start of each picture. When decoding the stream from any specific picture, it is possible to obtain SPS and PPS from the access unit including the picture and decode the picture accordingly.

  When cutting out the stream from any specific picture in video editing, it is guaranteed that SPS and PPS are included in the same access unit, so the video editing process becomes easy.

In video encoding, if the settings of encoding parameters are the same between pictures, the same parameter values are set for each picture, resulting in a redundant implementation. Figure 5 shows the example of placing SPS and PPS in each picture.

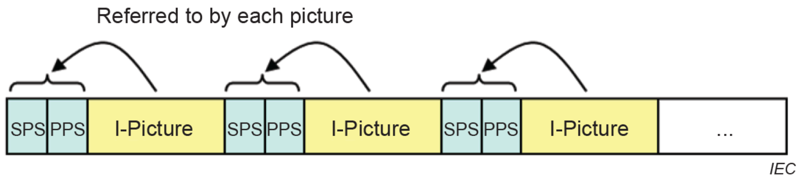

Figure 5 – Example of placing SPS and PPS in each picture

- In a long GOP stream, SPS and PPS necessary for decoding pictures in each GOP are placed in the access unit including an I Picture at the start of each GOP. It is common that the value of PPS is fixed for each picture type in the GOP.

  When decoding a long GOP stream, the decoder starts decoding from an I Picture at the start of the GOP. Since SPS and PPS can be obtained from the access unit at the start of the GOP, efficient decoding processing becomes possible.

  Simple video editing by cutting out the stream in the unit of the GOP, which includes the corresponding SPS and PPS, is also possible.

  In video encoding, even if encoding parameter settings are the same between GOPs, the same parameters need to be set for each GOP, resulting in a redundant implementation. In order to control image quality more finely in units of pictures, it is also possible to implement PPS in each picture unit in the GOP. Figure 6 shows the example of placing SPS and PPS at the start of each GOP.

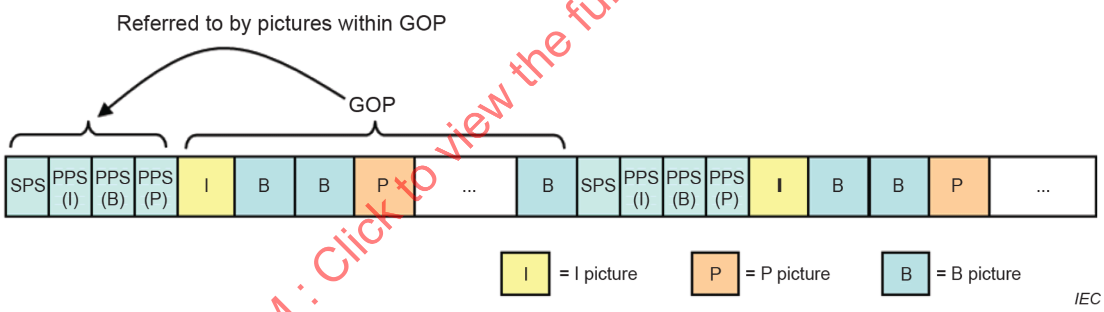

Figure 6 – Example of placing SPS and PPS at the start of each GOP

- If the values of SPS and PPS are constant throughout a sequence, SPS and PPS can be placed solely in the access unit at the beginning of the stream.

  Since adjusting the parameter values according to the characteristics of the picture is not performed at the time of encoding, the encoding process becomes simple. While this is advantageous in terms of bit rate because redundant SPS and PPS are not implemented, it is not possible to change the setting value of PPS according to the picture, which might be disadvantageous from the point of view of image quality.

  When decoding the stream from any specific picture, it is necessary to start the decoding process after acquiring SPS and PPS from the beginning of the stream. When segmenting the stream by editing, it is necessary to add SPS and PPS to the beginning of the clipped stream. Figure 7 shows the example of placing SPS and PPS at the start of the stream.

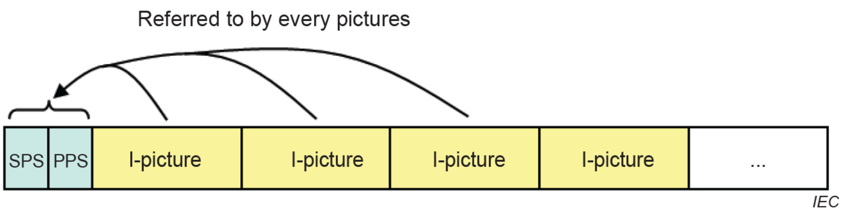

Figure 7 – Example of placing SPS and PPS at the start of the stream

5.7 Multi-slice encoding of AVC Stream

As described in 4.1, the blocks constituting the slice are the basic unit of codec processing. When decoding the stream, decoding results of adjacent blocks is required in order to perform various kinds of prediction calculation. Therefore, basically it is necessary to process each block sequentially. On the other hand, since the decode processing of each slice in one picture is basically independent, except for the de-blocking filter processing, which is performed in the final stage of picture decoding, it is possible to proceed in parallel.

Since modern computers tend to have multi-core processors, it is possible to achieve high performance with parallel processing in both hardware and software. Therefore, multi-slice encoding, which encodes one picture by dividing it into multiple slices, is recommended so that high-speed video decoding can be performed by using a multi-core-processor-based computer. Specifically, 1 to 8 slices are recommended in the case of a full high-definition picture and 4 to 16 slices are recommended in the case of a quad full high-definition (4K) picture.

However, considering the encoding process, it might be disadvantageous to divide the picture into multiple slices for encoding. This is because it is necessary to take into consideration the balance of compression ratio of the other slices in order to determine the optimal compression ratio of each slice. Therefore, cooperative processing on each slice is required.

Video compression utilizes the redundancy of the moving picture, but the redundancy of the video can vary considerably according to the picture content, and there can be extreme spatial deviation within a single picture. In other words, when dividing into multiple slices, the redundancy of the images contained in each slice could vary significantly.

Generally, when image redundancy is high, deterioration of image quality is not noticeable even if higher compression is used with a consequent decrease in data size, whereas if image redundancy is low, unless lower compression is used with a consequent increase in data size, deterioration of image quality could be highly noticeable. Therefore, when dividing into multiple slices, it is necessary to adaptively allocate the compressed data size to each slice. If the adaptive allocation process is difficult, there is a possibility that the image quality deteriorates locally. A single slice enables adaptive allocation as a series of processes within the slice, which is advantageous for image quality control.

In this way, it is necessary to decide whether or not to perform multi-slice encoding by considering advantages and disadvantages from the viewpoint of the entire system.