

团 体 标 准

T/AI 109.6—2022

信息技术 智能媒体编码 第 6 部分：智能媒体格式

Information technology - Intelligent media coding -

Part 6: Intelligent media format

2022 - 12 - 30 发布

2022 - 12 - 30 实施

中关村视听产业技术创新联盟 发布

目 次

前言.....II

引言.....III

1 范围.....1

2 规范性引用文件.....1

3 术语和定义.....2

4 缩略语.....4

5 AVS3 视频文件格式.....6

6 AVS3 视频 CMAF 轨道和媒体配置.....14

7 DASH 传输信令.....17

8 SMT 传输信令.....20

9 传输流和节目流技术要求.....26

附录 A（规范性） MIME 类型的‘codecs’参数.....34

附录 B（资料性） 知识位流编码媒体数据传输.....35

附录 C（资料性） AVS3 视频基本流的传输流和节目流封装和解封装.....39

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件是T/AI 109《信息技术 智能媒体编码》的第6部分。T/AI 109已经发布了以下部分：

——第2部分：视频；

——第4部分：视频符合性测试；

——第6部分：智能媒体格式。

本文件由数字音视频编解码技术标准（AVS）工作组提出。

本文件由中关村视听产业技术创新联盟归口。

本文件起草单位：中兴通讯股份有限公司、腾讯科技（深圳）有限公司、浙江大学、北京大学、上海交通大学、北京字节跳动网络技术有限公司、北京工业大学、上海大学、广东博华超高清创新中心有限公司。

本文件主要起草人：黄成、胡颖、虞露、于化龙、郑建铎、马思伟、徐异凌、白雅贤、李秋婷、许晓中、刘杉、王业奎、牟伦田、赵海武、袁錡超、林翔宇、张嘉琪、侯朴玥、侯礼志、殷骄阳、黄铁军、高文、张伟民、赵海英、龙仕强。

引 言

《T/AI 109 智能媒体编码》旨在确立智能媒体压缩的方法，拟由六个部分构成。

- 第 1 部分：系统。目的在于规定沉浸媒体系统中的虚拟现实视频和音频数据的封装格式、传输格式和信令格式。
- 第 2 部分：视频。目的在于规定了适应多种位率、分辨率和质量要求的高效视频压缩方法的解码过程。
- 第 3 部分：音频。目的在于描述面向全场景的下一代沉浸式音频高效压缩方法的解码过程。
- 第 4 部分：视频符合性测试。目的在于定义测试验证编码器生成的位流和解码器是否符合《信息技术 智能媒体编码 第 2 部分：视频》和《信息技术 智能媒体编码 第 3 部分：音频》所规定要求的方法。
- 第 5 部分：参考软件。目的在于定义满足《信息技术 智能媒体编码 第 2 部分：视频》和《信息技术 智能媒体编码 第 3 部分：音频》规定要求的参考软件。
- 第 6 部分：智能媒体格式。目的在于规定智能媒体编码数据的 AVS3 视频文件格式、AVS3 视频 CMAF 轨道和媒体配置、DASH 传输信令、SMT 传输信令、传输流和节目流技术要求。

本文件的发布机构提请注意，声明符合本文件时，可能涉及如下8项与数字视频编解码技术相关的专利的使用。专利申请号及名称如下：

序号	专利申请号	专利名称
1	CN202111309701.X	视频位流封装、解码、访问方法及装置
2	CN2018114887790	处理和传输媒体数据的方法和装置
3	CN2018114875469	指定参考图像的方法及装置及处理参考图像请求的方法及装置
4	CN202111125693.3	传输流和节目流的生成和及其装置
5	CN2019100028071	视频编解码方法与装置
6	CN201910888383.3	视频解码方法、视频编码方法、装置、设备及存储介质
7	CN201910503066.5	视频解码方法、视频编码方法、装置、设备及存储介质
8	CN201810974952.1	支持大跨度相关性信息编码的标识方法及系统

本文件的发布机构对上述专利的真实性、有效性和范围无任何立场。

上述专利持有人已向本文件的发布机构保证，愿意同任何申请人在合理且无歧视的条款和条件下，就专利授权许可进行谈判。上述专利持有人的声明已在本文件的发布机构备案，相关信息可以通过以下联系方式获得：

联 系 人：黄铁军（数字音视频编解码技术标准工作组秘书长）

通讯地址：北京大学理科2号楼2641室

邮政编码：100871

电子邮件：tjhuang@pku.edu.cn

电话：+8610-62756172

传真：+8610-62751638

网址：<http://www.avs.org.cn>

请注意除上述专利外，本文件的某些内容仍可能涉及专利。本文件的发布机构不承担识别专利的责任。

信息技术 智能媒体编码 第6部分：智能媒体格式

1 范围

本文件规定了智能媒体编码数据的 AVS3 视频文件格式、AVS3 视频 CMAF 轨道和媒体配置、DASH 传输信令、SMT 传输信令、传输流和节目流技术要求。

本文件适用于智能媒体编码系统中的视频直播、点播、网络流媒体等应用。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GY/T 155-2000 高清晰度电视节目制作及交换用视频参数值

GY/T 307-2017 超高清晰度电视系统节目制作和交换参数值

T/AI 109.2-2021 信息技术 智能媒体编码 第2部分：视频

T/AI 114-2021 信息技术 高效多媒体编码 第6部分：智能媒体传输

ISO/IEC 13818-1:2022 信息技术 运动图像及其伴音信息的通用编码 第1部分：系统 (Information technology -- Generic coding of moving pictures and associated audio information -- Part 1: Systems)

ISO/IEC 14496-12:2022 信息技术 音视频对象的编码 第12部分：ISO 基本媒体文件格式 (Information technology – Coding of audio-visual objects – Part 12: ISO base media file format)

ISO/IEC 14496-15:2021 信息技术 音视频对象的编码 第15部分：ISO 基本媒体文件格式承载网络抽象层单元结构化视频 (Information technology – Coding of audio-visual objects – Part 15: Carriage of network abstraction layer (NAL) unit structured video in the ISO base media file format)

ISO/IEC 23000-19:2020 信息技术 多媒体应用格式 第19部分：片段媒体通用媒体应用格式 (Information technology — Multimedia application format (MPEG-A) — Part 19: Common media application format (CMAF) for segmented media)

ISO/IEC 23009-1 信息技术 基于 HTTP 的动态自适应流媒体 第1部分：媒体呈现描述和片段格式 (Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 1: Media presentation description and segment formats)

IETF RFC 1738 统一资源定位符 (Uniform Resource Locators (URL))

IETF RFC 6381 "Bucket"媒体类型'Codecs'和'Profiles'参数 (The 'Codecs' and 'Profiles' Parameters for "Bucket" Media Types)

W3C XML 可扩展置标语言 (Extensible Markup Language (XML))

W3C XML Schema Part 1 可扩展置标语言模式定义语言 第一部分：结构 (XML Schema Definition Language (XSD) Part 1: Structures)

W3C XML Schema Part 2 可扩展置标语言模式定义语言 第二部分：数据类型 (XML Schema Definition Language (XSD) Part 2: Datatypes)

3 术语和定义

下列术语和定义适用于本文件。

3.1

AVS3视频编码位流 AVS3 video bitstream

符合T/AI 109.2-2021的编码图像所形成的二进制数据流。

3.2

AVS3视频存取单元 AVS3 video access unit

AVS3视频呈现单元的编码表达。

3.3

编码图像 coded picture

T/AI 109.2-2021中规定的一幅图像的编码表示。

3.4

表示 representation

封装有一个或多个具有描述性元数据的媒体成分（编码的音频、视频等）的结构化数据集合。

3.5

参考图像 reference picture

T/AI 109.2-2021中规定的解码过程中用于后续图像帧间预测的图像。

3.6

初始化片段 initialization segment

包含有媒体流解码所必需元数据的片段。

3.7

档次 profile

T/AI 109.2-2021中规定的语法、语义及算法的子集。

3.8

独立主位流 independent sequence stream

T/AI 109.2-2021 中规定的在不参考该位流以外的信息提供的知识图像的情况下可以完全解码的位流。

3.9

非独立主位流 **dependent sequence stream**

T/AI 109.2-2021 中规定的使用了该位流以外的知识位流提供的知识图像才能解码的位流。

3.10

轨道 **track**

ISO/IEC 14496-12:2022 标准中一系列相关样本的集合。

3.11

级别 **level**

T/AI 109.2-2021中规定的在某一档次下对语法元素和语法元素参数值的限定集合。

3.12

媒体呈现描述 **media presentation description**

用于提供流媒体服务的规范化描述媒体呈现的文件。

3.13

媒体片段 **media segment**

符合一定的媒体格式、可播放的片段。播放时可能需要与其前面的0个或多个片段以及初始化片段配合。

3.14

媒体资源 **asset**

任何与唯一标识符联系的用作构建一个多媒体演示的多媒体数据实体。

3.15

片 **patch**

T/AI 109.2-2021中规定的按光栅扫描顺序排列的相邻若干最大编码单元。

3.16

片段 **segment**

媒体呈现描述中的HTTP统一资源定位符引用的媒体单元。

3.17

RL图像 **RL picture**

T/AI 109.2-2021中规定的只使用知识图像作为参考图像进行帧间预测解码的P图像或B图像。

3.18

视频序列 **sequence**

T/AI 109.2-2021中规定的编码位流的最高层语法结构，包括一个或多个连续的编码图像。

3.19

随机访问 random access

T/AI 109.2-2021中规定的从某一点而非位流起始点开始对位流解码并恢复出解码图像的能力。

3.20

位流 bitstream

编码图像所形成的二进制数据流。

3.21

显示顺序 display order

显示解码图像的顺序。

3.22

样本 sample

在非提示轨道中，一个样本是一个单独的视频帧，时间连续的一个视频帧序列，或者时间连续的一段压缩音频。在提示轨道中，一个样本定义了一个或多个流式分组的构成。

3.23

知识图像 library picture

T/AI 109.2-2021中规定的知识位流中的图像，可以被其他位流中的图像参考。

3.24

知识位流 library stream

T/AI 109.2-2021 中规定的包含知识图像的位流。

3.25

主位流 sequence stream

T/AI 109.2-2021 中规定的可参考由该位流以外的信息提供的知识图像进行解码的位流。

3.26

子片段 subsegment

将片段按照时间间隔划分得到的部分片段（属于有效片段）。

3.27

自适应集 adaptation set

同一媒体内容的多个可替换的编码版本的集合。

4 缩略语

下列缩略语适用于本文件。

AVS	音视频编解码标准	(audio and video coding standard)
AVS3	第三代音视频编解码标准	(the third generation AVS standard)

CMAF	通用媒体格式	(common media application format)
CPB	编码图像缓冲区	(coded picture buffer)
DASH	基于 HTTP 的动态自适应流媒体	(dynamic adaptive streaming over HTTP)
DPB	解码图像缓冲区	(decoded picture buffer)
DTS	解码时间戳	(decoding time-stamp)
FIFO	先进先出	(first input first output)
HLG	混合对数型伽玛	(hybrid log gamma)
HTTP	超文本传输协议	(hypertext transfer protocol)
ISO BMFF	ISO 基本媒体文件格式	(ISO base media file format)
LBM	知识层数据缓存模型	(library buffer model)
LFU	最不常使用	(least frequently used)
LRU	最近最少使用	(least recently used)
LRAP	知识随机访问点	(library random access point)
MIME	多用途互联网邮件扩展类型	(multipurpose internet mail extensions)
MFU	媒体最小分割单元	(media fragment unit)
MP	媒体呈现	(media presentation)
MPD	媒体呈现描述	(media presentation description)
MTU	最大传输单元	(maximum transmission unit)
PA	包访问	(package access)
PAT	节目关联表	(program association table)
PQ	感知量化	(perceptual quantisation)
PES	分组化基本流	(packetized elementary stream)
PMT	节目映射表	(program map table)
PS	节目流	(program stream)
PSI	节目特定信息	(program-specific information)
PSM	节目流映射	(program stream map)
PTS	显示时间戳	(presentation time-stamp)
PID	包标识符	(packet identifier)
RAP	随机访问点	(random access point)
RTP	实时传输协议	(real-time transport protocol)
SAP	流访问点	(stream access point)
SDR	标准动态范围	(standard dynamic range)
SMT	智能媒体传输	(smart media transport)
SMTP	智能媒体传输协议	(SMT protocol)
STD	系统目标解码器	(system target decoder)
TCP	传输控制协议	(transmission control protocol)
TS	传输流	(transport kstream)
T-STD	传输系统目标解码器	(transport system target decoder)
UDP	用户数据报协议	(user datagram protocol)

URI	统一资源标识符	(uniform resource identifier)
URL	统一资源定位器	(uniform resource locator)
URN	统一资源名称	(uniform resource name)
UTC	协调世界时	(coordinated universal time)
UUID	通用唯一标识符	(universally unique identifier)
XML	可扩展置标语言	(extensible mark-up language)

5 AVS3 视频文件格式

5.1 AVS3 视频基本流和轨道定义

5.1.1 概述

T/AI 109.2-2021规定了适用于各种位率、分辨率和质量要求的高效视频压缩方法的解码过程。
AVS3视频编码位流的文件格式基于ISO/IEC 14496-12:2022 ISO基本媒体文件格式扩展。

5.1.2 AVS3 视频编码特性

AVS3视频编码位流的文件格式使用并扩展ISO/IEC 14496-12:2022规定的ISO基本媒体文件格式的
现有功能，支持下列AVS3视频编码特性：

——知识图像：知识图像是AVS3视频编码中用于帧间预测的特殊参考帧，RL图像是AVS3
视频编码中只使用知识图像作为参考图像进行帧间预测解码的图像。采用知识图像进行帧间
预测，可在特定应用场景中大幅提升编码效率；

——时域分层：AVS3视频编码支持将不同的编码图像按照解码时的依赖关系划分为不同
时域层级，划分为低层级的编码图像，在解码时无需参考更高层级的编码图像。

用于支持AVS3视频内容的结构化分组机制包括：

- 知识随机访问点样本群组：用于指示AVS3视频位流中作为LRAP样本的RL图像；
- 知识图像样本群组：用于指示AVS3视频位流中知识图像；
- 时域层级样本群组：用于指示与AVS3视频编码图像关联的时域可伸缩性的不同层级。

5.1.3 基本流结构

AVS3视频编码位流包括两种基本流的类型：

- 主基本流：可参考由该基本流以外的信息提供的知识图像进行解码的基本流；
- 知识基本流：只包含知识图像的基本流。其中，知识图像可以被其他基本流中的图像
参考，而不适用于显示输出。

使用上述基本流类型，存储AVS3视频内容的基本流中应包含序列头，并符合如下约束：

- 序列头应是基本流中包含的AVS3样本（同步样本或者知识随机访问点样本）的一部
分；
- 序列头也可以存储在AVS3视频轨道的样本入口中。

5.1.4 AVS3 视频轨道类型

用于封装AVS3视频基本流的AVS3视频轨道类型包括：

a) AVS3轨道:

- 1) AVS3轨道通过在样本入口中包含一个或多个序列头,并在其样本中封装AVS3视频编码图像来表示AVS3视频编码位流中的主基本流以及可能与其关联的知识基本流;
- 2) AVS3轨道通过在样本入口中包含一个或多个序列头,并在其样本中封装AVS3视频编码图像中的非知识图像来表示AVS3视频编码位流中的主基本流。此时,AVS3轨道可以引用AVS3知识位流轨道;
- 3) 当AVS3轨道通过'a31r'轨道引用引用AVS3知识位流轨道时,也称为AVS3主位流轨道;

b) AVS3知识位流轨道: AVS3知识位流轨道通过在样本入口中包含一个序列头,并在其样本中封装AVS3视频编码图像中的知识图像来表示AVS3视频编码位流中的知识基本流。AVS3知识位流轨道由AVS3主位流轨道通过'a31r'轨道引用来引用。

5.1.5 时域分层存储

AVS3视频位流的时域层级支持多轨道存储。多轨道存储可用于选择性地访问和播放AVS3视频位流的时域层级。

如果AVS3视频编码位流支持时域分层,即序列头的“时间层标识允许标志字段”取值‘1’,AVS3视频编码位流中可以包含一个或多个不同域层级的编码图像。AVS3视频主基本流中不同域层级的编码图像可通过以下方式存储:

a) 所有域层级的编码图像存储在同一个AVS3主位流轨道:

- 1) AVS3主位流轨道应使用5.3.1中定义的样本入口类型'av3';
- 2) AVS3主位流轨道应使用5.4.3中定义的时域层级样本群组来为轨道中的样本提供时域层级信息;

b) 所有域层级的编码图像存储在多个不同的AVS3主位流轨道。其中,每个AVS3轨道中包含一个或多个时域层级对应的编码图像:

- 1) 包含最低域层级的AVS3主位流轨道应使用5.3.1中定义的样本入口类型'av3';
- 2) 不包含最低域层级的AVS3主位流轨道应使用5.3.2中定义的样本入口类型'lav3';
- 3) 其中,样本入口类型为'lav3'的AVS3轨道应使用'ltrf'类型的轨道引用,引用样本入口类型为'av3'的AVS3轨道,并且可选地引用一个或多个样本入口类型为'lav3'的包含较低域层级的AVS3轨道。通过解析'ltrf'类型的轨道引用,包含不同域层级的一个或多个AVS3轨道中的所有样本应包含解码至所有域层级中的最高域层级对应的主基本流中编码图像子集所需的所有数据。

AVS3时域分层位流重建流程示例如下:

示例:

当AVS3视频编码位流中不同域层级的编码图像存储在不同的AVS3轨道时,解码AVS3视频编码位流中时域分层的编码图像之前,客户端需要获取包含对应域层级的所有AVS3轨道,重建AVS3视频编码位流中对应域层级的子位流。

客户端首先确定AVS3视频编码位流中所要解码的最高域层级,并获取包含该域层级的AVS3轨道。客户端通过解析'ltrf'类型的轨道引用,获取包含解码至所述最高域层级所需AVS3视频编码位流子集的一个或多个AVS3轨道。

客户端按照解码时间对所获取的一个或多个AVS3轨道中的所有编码图像进行重新排序,重建AVS3视频编码位流,并按照5.2.2中规定的解码器配置信息中的约束进行解码。

5.2 样本与解码器配置定义

5.2.1 样本格式

AVS3样本：AVS3轨道中的样本应包含T/AI 109.2-2021中规定的一个编码图像，还可包含T/AI 109.2-2021中规定的序列头和序列结束码。

AVS3知识图像样本：AVS3知识位流轨道中的样本包含T/AI 109.2-2021中规定的一个知识图像，还可包含T/AI 109.2-2021中规定的序列头和序列结束码。

5.2.2 解码器配置信息

5.2.2.1 AVS3 解码器配置记录

5.2.2.1.1 定义

AVS3解码器配置记录适用于不支持时域分层或者包含最低时域层级的AVS3视频位流。

5.2.2.1.2 语法

```
aligned(8) class Avs3DecoderConfigurationRecord {
    unsigned int(8) configurationVersion = 1;
    unsigned int(16) sequence_header_length;
    bit(8*sequence_header_length) sequence_header;
    bit(6) reserved = '111111'b;
    unsigned int(2) library_dependency_idc;
}
```

5.2.2.1.3 语义

sequence_header_length：以字节为单位指示配置记录所适用的基本流中序列头的长度。

sequence_header：包含 T/AI 109.2-2021 中定义的 sequence_header()，并符合以下约束：

- 如果配置记录适用于主基本流，则序列头中的library_stream_flag字段取值‘0’。而如果配置记录适用于知识基本流，则序列头中的library_stream_flag字段取值‘1’；
- 如果配置记录适用于主基本流，并且该主基本流不参考知识基本流中知识图像进行解码，则序列头中 library_picture_enable_flag 字段取值‘0’。如果配置记录适用于主基本流，并且该主基本流参考知识基本流中的知识图像进行解码，则序列头中 library_picture_enable_flag 字段取值‘1’。

library_dependency_idc：指示配置记录所适用的基本流的类型。如果配置记录适用于主基本流，library_dependency_idc 还指示主基本流是否参考知识基本流中的知识图像进行解码。library_dependency_idc 的值与描述见表1。

表1 Library_dependency_idc 的值与描述

值	描述
00	当前解码器配置记录适用于主基本流，且该主基本流不参考知识基本流中的知识图像进行解码
01	当前解码器配置记录适用于主基本流，且该主基本流参考知识基本流中的知识图像进行解码
10	当前解码器配置记录适用于知识基本流
11	当前解码器配置记录所对应的轨道中同时封装主基本流和知识基本流，且主基本流参考知识基本流中的知识图像进行解码

5.2.2.2 LAVS3 解码器配置记录

5.2.2.2.1 定义

LAVS3解码器配置记录适用于需要参考一个或多个较低时域层级进行解码的AVS3视频位流，AVS3视频位流存储在多个不同AVS3轨道中，每个轨道都包含一个或多个高层时域层级。

5.2.2.2.2 语法

```
aligned(8) class LAvs3DecoderConfigurationRecord {
    unsigned int(8) configurationVersion = 1;
    unsigned int(8) num_temporal_layers;
    for(i=0; i< num_temporal_layers; i++) {
        unsigned int(3) temporal_layer_id[i];
        unsigned int(4) frame_rate_code[i];
        bit(6) reserved = '1'b;
        unsigned int(18) temporal_bit_rate_lower[i];
        unsigned int(12) temporal_bit_rate_upper[i];
        bit(2) reserved = '11'b;
    }
}
```

5.2.2.2.3 语义

num_temporal_layers: 指示解码器配置记录所适用的基本流中包含的时域层级数目。

temporal_layer_id[i]: 指示解码器配置记录所适用的基本流中包含的时域层级标识。

frame_rate_code[i]: 指示当时域层级累计到 temporal_layer_id[i]时，AVS3 视频编码位流对应的帧率代码。

temporal_bit_rate_lower[i]: 指示当时域层级累计到 temporal_layer_id[i]时，AVS3 视频编码位流对应码率的低 18 位。

temporal_bit_rate_upper[i]: 指示当时域层级累计到 temporal_layer_id[i]时，AVS3 视频编码子位流对应码率的高 12 位。

5.2.2.3 AVS3 解码器配置数据盒

5.2.2.3.1 定义

Avs3ConfigurationBox包含5.2.2.1.1规定的AVS3解码器配置记录。

5.2.2.3.2 语法

```
class Avs3ConfigurationBox extends Box('av3c',version = 0) {  
    Avs3DecoderConfigurationRecord () Avs3Config;  
}
```

5.2.2.3.3 语义

Avs3DecoderConfigurationRecord在5.2.2.1.1中定义。

5.2.2.4 LAVS3 解码器配置数据盒

5.2.2.4.1 定义

LAvs3ConfigurationBox包含5.2.2.1规定的LAVS3解码器配置记录。

5.2.2.4.2 语法

```
class LAvs3ConfigurationBox extends Box('lavc',version = 0) {  
    LAvs3DecoderConfigurationRecord () LAvs3Config;  
}
```

5.2.2.4.3 语义

LAvs3DecoderConfigurationRecord在5.2.2.1中定义。

5.3 ISO 基本媒体文件格式扩展

5.3.1 AVS3 样本入口

5.3.1.1 定义

样本入口类型: 'avs3'

容器: Sample Table Box ('stbl')

强制性: 封装 AVS3 视频位流的轨道中的至少一个轨道应有一个 'avs3' 样本入口

数量: 一个或多个

AVS3轨道的每个样本入口应是一个AVS3样本入口，其样本入口类型等于 'avs3' 或者5.3.2.1定义的 'lav3'。AVS3轨道应符合如下约束：

- 包含符合T/AI 109.2-2021规定的AVS3视频位流的AVS3轨道的每个样本入口应包含一个Avs3ConfigurationBox;
- 包含最低时域层级的AVS3轨道的每个样本入口应包含一个Avs3ConfigurationBox，以及一个LAvs3ConfigurationBox;
- 不包含最低时域层级的AVS3轨道的每个样本入口应包含一个LAvs3ConfigurationBox。

5.3.1.2 语法

```
class Avs3SampleEntry() extends VisualSampleEntry ('avs3') {
    Avs3ConfigurationBox  config;
    LAvs3ConfigurationBox (); // optional
    MPEG4ExtensionDescriptorsBox (); // optional
}
```

5.3.1.3 语义

Avs3ConfigurationBox: 提供 AVS3 视频编码位流的解码配置信息。

LAvs3ConfigurationBox: 提供支持时域分层的 AVS3 视频编码位流的解码配置信息。

MPEG4ExtensionDescriptorsBox: 定义见 ISO/IEC 14496-15:2021 中 5.4.2。

Compressorname: 基类 VisualSampleEntry 中指示使用的编码器的名称，建议使用值“\013AVS3 Coding”（\013 代表数字 11，表示以字节为单位字符串的长度）。

5.3.2 时域分层样本入口

5.3.2.1 定义

样本入口类型: 'lav3'

容器: Sample Table Box ('stbl')

强制性: 'lav3' 样本入口是强制的

数量: 一个或多个

当 AVS3 轨道不包含符合 T/AI 109.2-2021 中规定的 AVS3 视频位流时，应使用样本入口类型等于 'lav3' 的 LAvs3SampleEntry。

5.3.2.2 语法

```
class LAvs3SampleEntry() extends VisualSampleEntry ('lav3') {
    LAvs3ConfigurationBox ();
    MPEG4ExtensionDescriptorsBox (); // optional
}
```

5.3.3 受限样本入口

AVS3 知识位流轨道仅包含用于参考而不用于显示输出的 AVS3 知识图像样本。

AVS3 知识位流轨道在文件中表示为受限视频，应使用符合下列附加约束的受限样本入口 'resv':

- 原始样本入口类型应包含在 RestrictedSchemeInfoBox 中的 OriginalFormatBox。其中，OriginalFormatBox 的 data_format 字段取值 'avs3'；
- SchemeTypeBox 存在于 RestrictedSchemeInfoBox 中，并且 scheme_type 设置为 'av31'。

5.3.4 同步样本

AVS3 视频轨道（样本入口类型为 'avs3'）中包含随机访问点编码图像的样本被定义为同步样本，同步样本的使用符合 ISO/IEC 14496-12:2022 中 8.6.2 的定义。

当样本入口类型为'avs3'时，适用以下约束：

- a) 如果样本是同步样本或者知识随机访问点样本，解码该样本所需的序列头应包含在该样本中。此外，解码该样本所需的序列头还可以包含在样本入口中；
- b) 否则（样本不是同步样本或者知识随机访问点样本），解码样本所需的序列头应包含在该样本的前一个同步样本或知识随机访问点样本中。此外，解码样本所需的序列头还可以包含在样本入口中。

5.3.5 LRAP 样本

AVS3主位流轨道中的RL图像样本只使用AVS3知识位流轨道中的知识图像样本作为参考图像进行帧间预测解码。

AVS3主位流轨道中符合以下特性的RL图像样本被定义为LRAP样本，即，如果作为RL图像的参考图像的所有知识图像均可以获得，则按照解码顺序和显示顺序在LRAP图像之后的所有图像都可以正确解码。

LRAP样本具有5.4.1中定义的知识随机访问点类型（LRAP_type）。

5.3.6 片段索引数据盒

5.3.6.1 定义

数据盒类型： 'lidx'

容器： File

强制性： 否

数量： 零个或多个

LRAPSegmentIndexBox指示其引用的子片段中存在5.3.5中定义的知识随机访问点，并符合：

- a) 如果子片段封装的轨道分段包含LRAP，则子片段包含LRAP，并以LRAP开始；
- b) 子片段中的LRAP样本（LRAP_type类型取值‘1’、‘2’或‘3’）应被标记为群组类型为'lrap'的样本群组的成员。

5.3.6.2 语法

```
aligned(8) class LRAPSegmentIndexBox extends SegmentIndexBox('lidx',
version, 0) {
    unsigned int(16) reserved = 0;
    unsigned int(16) reference_count;
    for(i=1; i <= reference_count; i++) {
        unsigned int(1) starts_with_LRAP;
        unsigned int(3) LRAP_type;
        unsigned int(28) reserved = 0;
    }
}
```

5.3.6.3 语义

reference_count： 提供被引用项目的数量。

`starts_with_LRAP`: 取值为‘1’表示引用的子片段以 **LRAP** 开始; `starts_with_LRAP` 取值为‘0’表示引用的子片段不以 **LRAP** 开始。

`LRAP_type`: 指示子片段中 **LRAP** 的知识随机访问点类型。

5.4 样本群组定义

5.4.1 LRAP 样本群组

5.4.1.1 定义

LRAP样本群组指示视频位流中的**LRAP**样本, 以及作为参考图像样本的知识图像样本的信息。**LRAP**样本群组的成员应符合以下条件:

- LRAP**样本只能以知识图像样本作为参考图像样本;
- 在解码其所参考的知识图像样本之后, 从**LRAP**样本处开始解码可以正确解码**LRAP**样本以及所有显示顺序在该**LRAP**样本之后的样本。

注: **LRAP**样本只能与其所参考的知识图像样本联合使用。为创建可解码的样本序列, 需要将知识图像样本与**LRAP**样本以及按显示顺序在**LRAP**样本之后的样本串联起来。

5.4.1.2 语法

```
class VisualLRAPEntry extend VisualSampleGroupEntry ('lrap') {
    unsigned int(3) LRAP_type;
    unsigned int(3) entry_count;
    unsigned int(2) reserved = 0;
    int i;
    for (i=0; i < entry_count; i++) {
        unsigned int(9) library_sample_number;
        unsigned int(7) reserved = 0;
    }
}
```

5.4.1.3 语义

`LRAP_type` 是一个非负整数, 表示假设 **LRAP** 样本不依赖于知识图像样本时, **LRAP** 样本所对应的 ISO/IEC 14496-12:2022 附录 I 中定义的流访问点类型 (`SAP_type`)。 `LRAP_type` 在 1 到 3 的范围内取值, 其他类型值保留。

`entry_count` 是一个非负整数, 表示群组中 **LRAP** 样本参考的知识图像样本的数量。`entry_count` 值为‘001’表示群组中 **LRAP** 样本参考的知识图像样本数为 1; 值‘000’和‘010’‘111’保留。

`library_sample_number` 是一个非负整数, 提供群组中 **LRAP** 样本所参考的知识图像样本的样本编号。

5.4.2 知识图像样本群组

5.4.2.1 定义

当AVS3轨道中包含来自自主基本流和知识基本流的样本时，知识图像样本群组（'a3lg'）用于标识AVS3轨道中的知识图像样本。

5.4.2.2 语法

```
class LibrarySampleGroupEntry extends VisualSampleGroupEntry ('a3lg') {
}
```

5.4.3 时域层级样本群组

5.4.3.1 定义

时域层级样本群组（'telg'）可用于根据时域层级对AVS3轨道中的样本进行分组，并为组中的样本提供时域层级信息。

5.4.3.2 语法

```
class TemporalLayerEntry extend VisualSampleGroupEntry ('telg') {
    unsigned int(8) temporal_layer_id;
}
```

5.4.3.3 语义

temporal_layer_id 指示样本群组中的样本包含的编码图像的时域层级标识。

6 AVS3 视频 CMAF 轨道和媒体配置

6.1 概述

AVS3 视频 CMAF 轨道格式应符合 ISO/IEC 23000-19:2020 中 9.2 指定的通用视频 CMAF 轨道格式、第 5 章中指定的 AVS3 视频轨道格式和第 6 章中指定的约束。

如果CMAF轨道符合上述要求，则它被称为AVS3 视频CMAF轨道，并且其品牌标识定义为'ca3v'。

6.2 CMAF 轨道约束

6.2.1 基本要求

任何符合 AVS3 媒体配置文件的 CMAF 轨道都应符合 5.1.4 中定义的 AVS3 轨道、ISO/IEC 23000-19 中定义的通用视频 CMAF 轨道格式和第 6 章中定义的附加约束。

6.2.2 轨道头数据盒

对于AVS3视频CMAF轨道，轨道头数据盒中宽度和高度的值应归一化为AVS3视频的宽度和高度，定义如下：

- 归一化显示的高度应为垂直视频空间样本的数量，即序列头中的vertical_size；
- 归一化显示的宽度应为水平视频空间样本的数量，即序列头中的horizontal_size，乘以视频空间样本宽高比。其中，序列头中的aspect_ratio指定AVS3视频的样本宽高比；
- 不应存在CleanApertureBox数据盒，裁剪后的光圈被定义为视频CMAF轨道中的活动图像（'clean'）。

6.2.3 样本描述数据盒

AVS3 视频轨道中的 SampleDescriptionBox 应包含符合 ISO/IEC 14496-12:2022 中规定的一个或多个视觉样本入口。

符合 AVS3 视频媒体配置文件的 CMAF 轨道的视觉样本入口的语法和取值应符合 5.3 中定义的 AVS3SampleEntry('avs3') 或 LAVS3SampleEntry('lav3') 样本入口。

SampleDescriptionBox 中的第一个 VisualSampleEntry 符合：

- a) 应包含 width 和 height 字段，且其值等于或大于视频轨道中视频片所引用的任何序列头中的最大水平和垂直视频空间样本计数值；
- b) 包含解码器配置记录，且包含 CMAF 轨道中至少一个 CMAF 分段所需的编解码器“档次”值和“级别”值；
- c) 包含一个或多个 ISO/IEC 14496-12:2022 中规定的 ColourInformationBox，其 color_type 字段值等于'nclx'，且符合：
 - 1) colour_primaries 的高 8 位取值 '0'，低 8 位携带 T/AI 109.2-2021 中定义的彩色三基色 colour_primaries 字段值；
 - 2) transfer_characteristics 的高 8 位取值 '0'，低 8 位携带 T/AI 109.2-2021 中定义的光电转移特性 transfer_characteristics 字段值；
 - 3) matrix_coefficients 的高 8 位取值 '0'，低 8 位携带 T/AI 109.2-2021 中定义的彩色信号转换矩阵 matrix_coefficients 字段值；
 - 4) full_range_flag 携带 T/AI 109.2-2021 中定义的样值范围 sample_range 字段值。

6.2.4 AVS3 视频编码图像

依照 ISO/IEC 14496-12:2022 中定义，每个媒体样本应包含一个特定于一个呈现时间和持续时间的 AVS3 视频编码图像。

每个 AVS3 视频编码图像应作为媒体样本存储在 CMAF 块和/或 CMAF 分段中的 MediaDataBox 数据盒。

6.2.5 AVS3 视频编码序列

每个 CMAF 分段应包含一个或多个符合 T/AI 109.2-2021 所规定的完整编码视频序列。

每个 CMAF 分段中的 AVS3 视频样本描述应在编码视频序列的第一个编码图片中包含该编码视频序列引用的序列头语法结构。

所有 CMAF 分段中的第一个媒体样本应是下面指定的随机访问点之一：

- a) 每个 CMAF 分段中的第一个媒体样本是 ISO/IEC 14496-12:2022 定义的 SAP_type 为 '1' 或 '2' 的 SAP 样本。为此，知识图像不应作为参考图像出现在编码视频序列中，即，编码视频序列中的首个序列头符合以下约束：
 - 1) library_stream_flag 字段的值应设置为 '0'；
 - 2) library_picture_enable_flag 字段的值应设置为 '0'；
- b) 每个 CMAF 分段中第一个媒体样本是 5.4.1 中定义的 LRAP_type 为 '1' 或 '2' 的 LRAP 样本。知识图像可以作为参考图像出现在编码视频序列中，即，编码视频序列的首个序列头符合以下约束：

- 1) `library_stream_flag` 字段的值应设置为 ‘0’ ；
- 2) `library_picture_enable_flag` 字段的值应设置为 ‘1’ 。

注：LRAP 样本解码时间之前，作为参考图像样本的所有知识图像样本均应可以获得。

6.3 CMAF 切换集约束

6.3.1 基本要求

对于符合 AVS3 视频媒体配置文件的 CMAF 切换集，适用以下约束：

- a) CMAF 交换集中的每个 CMAF 轨道应符合 6.2 中定义的 CMAF 轨道约束；
- b) CMAF 切换集中的每个 CMAF 轨道应符合 ISO/IEC 23000-19:2020 中 7.3.4 规定的通用 CMAF 切换集约束和 ISO/IEC 23000-19:2020 中 9.2.11 规定的通用视频 CMAF 切换集要求；
- c) 单一初始化 CMAF 切换集应符合 6.3.2 中定义的附加约束。

6.3.2 单一初始化 CMAF 切换集约束

AVS3 视频 CMAF 切换集应符合如下指定的单一初始化约束：

- a) 应符合 ISO/IEC 23000-19:2020 中 9.2.11.4 规定的通用单一初始化约束；
- b) 应存储每个视频片所引用的序列头，用于解码和显示，规定如下：
 - 1) 带内存储应用于 'avs3' 样本描述 CMAF 轨道。序列头以视频序列起始码开始，后跟一串编码图片。序列头可以在位流中重复出现，称为重复序列头；
 - 2) 样本入口存储和带内存储的组合可用于 'avs3' 样本描述 CMAF 轨道。在这种情况下，任何 CMAF 头都应提供 CMAF 交换集中的任何 CMAF 分段引用的所有样本入口；
- c) 如果所有 CMAF 头部中都包含 CMAF 媒体配置品牌，每个 CMAF 头部中的第一个视觉样本入口应足以初始化 CMAF 切换集中的所有 CMAF 轨道或者符合 CMAF 媒体配置的 CMAF 轨道的解码、解密和显示；
- d) 如果 CMAF 切换集中的 CMAF 轨道不包含 CMAF 媒体配置品牌，则任何 CMAF 头部应足以初始化 CMAF 切换集一次并解码 CMAF 切换集中的所有 CMAF 分段；
- e) 如果 CMAF 切换集中的所有 CMAF 轨道包含 CMAF 媒体配置品牌，则具有特定 CMAF 媒体配置品牌的任何 CMAF 头部应足以初始化一次并解码具有相同 CMAF 媒体配置品牌，或符合初始化 CMAF 媒体配置品牌的其他 CMAF 轨道。

6.4 AVS3 视频基本流约束

6.4.1 基本要求

符合 AVS3 视频媒体配置的 CMAF 切换集的 CMAF 轨道中包含的 AVS3 视频基本流，应遵守 T/AI 109.2-2021 中定义的通用配置和级别语义。

6.4.2 序列头约束

符合 AVS3 视频媒体配置的 CMAF 轨道内的序列头应符合 T/AI 109.2-2021 的定义并具有以下附加约束：

- a) 以下字段应具有如下预定值：
 - 1) `progressive_sequence` 字段的值应等于 ‘1’ ；

- 2) `field_coded_sequence` 字段的值应等于 '0' ;
- b) 在一个 CMAF 轨道以及一个 CMAF 切换集中的任何 CMAF 轨道中, 以下每个字段的值不应从一个编码视频序列变为另一个:
 - 1) `profile_id`;
 - 2) `level_id`。

6.5 视频编解码参数

呈现应用程序应使用符合 IETF RFC 6381 中规定的参数发送 AVS3 视频 CMAF 轨道和 CMAF 切换集的视频编解码器配置和级别。

AVS3 视频媒体配置的 MIME 类型的“编解码器”参数见附录 A。

7 DASH 传输信令

7.1 概述

AVS3 视频编码位流支持通过 ISO/IEC 23009-1 基于 HTTP 的动态自适应流媒体传输协议进行传输, 用于封装 AVS3 视频编码位流的媒体呈现描述与片段格式基于 ISO/IEC 23009-1 扩展。

7.2 DASH 片段格式

7.2.1 基本要求

用于封装 AVS3 视频数据的 DASH 片段格式应符合第 5 章规定的 AVS3 视频编码位流的文件格式, 其样本入口类型应支持 'avs3'。

7.2.2 初始化片段

每个 DASH 初始化片段应包含一个 `Avs3ConfigurationBox`, 其中包括 `Avs3DecoderConfigurationRecord` 解码器配置记录。

7.2.3 媒体片段

每个 DASH 媒体片段应包含一个或多个 T/AI 109.2-2021 中规定的编码图像。

AVS3 视频编码图像解码所需的序列头信息应在媒体片段中获取。

每个 DASH 媒体片段中的第一个媒体样本应符合以下任意一个约束:

- a) DASH 媒体片段中的第一个媒体样本是 ISO/IEC 14496-12:2022 附录 I 定义的 `SAP_type` 为 1 或 2 的 SAP 样本。知识图像不应作为参考图像出现在编码视频序列中, 即, 编码视频序列中的首个序列头语法结构实例符合以下约束:
 - 1) `library_stream_flag=0`;
 - 2) `library_picture_enable_flag=0`;
- b) DASH 媒体片段中的第一个媒体样本是 `LRAP_type` 为 '1' 或 '2' 的 LRAP 样本。知识图像可以作为参考图像出现在编码视频序列中, 即, 编码视频序列中首个序列头语法结构实例符合以下约束:
 - 1) `library_stream_flag=0`;

2) `library_picture_enable_flag=1`;

注：LRAP 样本解码时间之前，作为参考图像样本的所有知识图像样本均应可以获得。

7.2.4 索引片段

索引片段提供动态自适应流媒体中的一个表示的所有媒体片段的索引。索引片段以一个'`styp`'类型数据盒开始，表示中每个媒体片段包含若干个连续子片段并由一个或多个符合以下约束条件之一的片段索引数据盒索引：

- a) 每个子片段由一个 ISO/IEC 14496-12:2022 中 8.16.3 定义的'`sidx`'类型的片段索引数据盒索引。'`sidx`'数据盒指示引用的子片段是否以 SAP 开头，以及子片段中流访问 SAP 的 `SAP_type` 类型；
- b) 每个子片段由一个 5.3.6 中定义的'`lidx`'类型的片段索引数据盒索引。'`lidx`'数据盒指示引用的子片段是否以 LRAP 开头，以及子片段中 LRAP 的 `LRAP_type` 类型。

7.3 DASH MPD 编码器参数

AVS3 视频编码位流在媒体呈现描述中使用 `@codecs` 属性表示“档次”和“级别”，属性值应符合 IETF RFC 6381 和附录 A。

7.4 DASH MPD 描述子

7.4.1 XML 命名空间和规则

本文件中定义并使用了一些新的 XML 元素和属性，这些元素和属性在命名空间“`urn:avs:avs3:p6:2022`”中定义，命名空间指示符“`avs3`”用于指代此命名空间（智能媒体格式）。命名空间指示符“`xs:`”的定义符合 XML Schema Part 1，数据类型及其含义符合 XML Schema Part 2。

7.4.2 时域层级描述子

一个 `@schemeIdUri` 属性为“`urn:avs:avs3:p6:2022:highest_temporal_id`”的补充属性元素为时域层级描述子，用于指示最高时域层级信息。时域层级描述子的 `@value` 属性值为最高时域层级标识值。

表示中可以包含多个不同时域层级的编码图像。当 AVS3 视频编码位流支持时域分层时，可以根据 AVS3 视频的主位流中不同时域层级的编码图像生成 DASH 的一个或多个表示，每个表示中应当包含一个时域层级描述子，并且每个表示中应包含从最低时域层级到时域层级描述子指示的最高时域层级的全部时域层级。表示中的 AVS3 视频序列头可以应用于当前表示的全部时域层级。

7.4.3 知识位流依赖描述子

一个 `@schemeIdUri` 属性为“`urn:avs:avs3:p6:2022:LibraryDependency`”的基本属性元素为知识位流依赖描述子。知识位流依赖描述子应包含于表示层级，用于指示描述子所适用的表示中封装的基本流类型以及知识位流依赖信息，其中基本流类型包括独立主位流、非独立主位流、知识位流。

该描述子的 `@value` 元素不应出现，知识位流依赖描述子包含的元素及属性见表 2。

表2 知识位流依赖描述子元素及属性

描述子元素及属性	类型	使用	描述
LibraryDependency	avs3:LibraryDependencyType	0..1	该元素指示表示层级的 AVS3 位流信息
LibraryDependency@library_dependency_idc	xs:unsignedByte	必选	取值为 0 时指示当前表示仅包含独立主位流； 取值为 1 时指示当前表示仅包含非独立主位流； 取值为 2 时指示当前表示仅包含知识位流； 取值为 3 时指示当前表示包含非独立主位流和知识位流
LibraryDependency.LibraryInfo	avs3:LibraryInfoType	0..N	@library_dependency_idc 属性取值为 1 或 3 时使用该元素指示表示中的媒体片段与包含知识图像的媒体片段间存在依赖关系，其属性用于描述被依赖的媒体片段的定位信息和知识图像索引
LibraryDependency.LibraryInfo@library_pid	xs:unsignedInt	必选	指示被依赖媒体片段中包含的知识图像在编解码的索引
LibraryDependency.LibraryInfo@segmentURL	xs:anyURL	必选	指示被依赖媒体片段的定位信息（URL 或 DASH MPD 中的媒体片段标识符）

LibraryDependency@library_dependency_idc 取值为 ‘1’ 的表示应通过 @associationId 关联至相应知识位流轨道对应的表示（该表示的 LibraryDependency@library_dependency_idc 取值为 ‘2’），且 @associationType 取值应等于 ‘a3lr’。

位流依赖描述子的 XML Schema 如下所示：

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
  targetNamespace="urn:avs:avs3:p6:2022"
  xmlns:avs3="urn:avs:avs3:p6:2022"
  elementFormDefault="qualified">
  <xs:element name="LibraryDependency" type="avs3:LibraryDependencyType"/>
  <xs:complexType name="LibraryDependencyType">
    <xs:sequence>
      <xs:element name="LibraryInfo">
        <xs:complexType>
          <xs:attribute name="library_pid" type="xs:integer" use="required" />
          <xs:attribute name="segmentURL" type="xs:anyURI" use="required" />
        </xs:complexType>
      </xs:sequence>
      <xs:attribute name="library_dependency_idc" type="xs:unsignedByte"
use="required" />
    </xs:complexType>
  </xs:schema>
```

7.4.4 彩色信息描述子

AVS3 视频编码位流使用的彩色三基色、彩色信号转换矩阵、光电转移特性信息通过基本属性元素描述子在 MPD 中进行描述，见表 3，AVS3 彩色信息描述子应包含于自适应集层级。

表3 AVS3 彩色信息描述子

@schemeIdUri	基本属性元素描述子
urn:avs:avs3:p6:2022:ColourPrimaries	表示 AVS3 视频编码位流使用的彩色三基色，其@value 取值应与媒体片段包含的 AVS3 视频编码位流中 colour_primaries 的值相同
urn:avs:avs3:p6:2022:MatrixCoefficients	表示 AVS3 视频编码位流使用的彩色信号转换矩阵，其@value 取值应与媒体片段包含的 AVS3 视频编码位流中 matrix_coefficients 的值相同
urn:avs:avs3:p6:2022:TransferCharacteristics	表示 AVS3 视频编码位流使用的光电转移特性，其@value 取值应与媒体片段包含的 AVS3 视频编码位流中 transfer_characteristics 的值相同

AVS3 彩色信息描述子应符合以下约束：

- a) 当自适应集中的 AVS3 视频编码位流支持 GY/T 155-2000 和 SDR 时，应包含以下基本属性元素描述子：
 - 1) @schemeIdUri=urn:avs:avs3:p6:2022:ColourPrimaries 且@value='1'；
 - 2) @schemeIdUri=urn:avs:avs3:p6:2022:MatrixCoefficients 且@value='1'；
 - 3) @schemeIdUri=urn:avs:avs3:p6:2022:TransferCharacteristics 且@value='1'或'6'；
- b) 当自适应集中的 AVS3 视频编码位流支持 GY/T 307-2017 和 HLG 时，应包含以下基本属性元素描述子：
 - 1) @schemeIdUri=urn:avs:avs3:p6:2022:ColourPrimaries 且@value='9'；
 - 2) @schemeIdUri=urn:avs:avs3:p6:2022:MatrixCoefficients 且@value='8'或'9'；
 - 3) @schemeIdUri=urn:avs:avs3:p6:2022:TransferCharacteristics 且@value='14'；
- c) 当自适应集中的 AVS3 视频编码位流支持 GY/T 307-2017 和 PQ 时，应包含以下基本属性元素描述子：
 - 1) @schemeIdUri=urn:avs:avs3:p6:2022:ColourPrimaries 且@value='9'；
 - 2) @schemeIdUri=urn:avs:avs3:p6:2022:MatrixCoefficients 且@value='8'或'9'；
 - 3) @schemeIdUri=urn:avs:avs3:p6:2022:TransferCharacteristics 且@value='12'。

8 SMT 传输信令

8.1 概述

AVS3 视频编码位流支持通过 T/AI 114-2021 SMT 协议进行传输，传输信令基于 T/AI 114-2021 扩展。

在 SMT 中，PA 消息用于指示媒体数据包消费，一个 PA 消息应当包含一个 PA 表、一个 MP 表和一个媒体呈现层信息表。其中，MP 表包含一个或多个描述符，用于指示媒体资源的媒体数据包信息，通过读取描述符的标识符 descriptor_tag 确定该描述符的类型。

8.2 缓存内容更新信令

8.2.1 LBM 消息

8.2.1.1 定义

LBM 消息由服务器发送给客户端，包含对非对齐时间段知识层数据的最佳存储大小、存储管理方法（例如 FIFO、LFU、LRU 等存储管理方法）等信息。

8.2.1.2 语法

LBM消息的语法见表4。

表4 LBM 消息语法

语法	值	位数	类型
LBM_message () { message_id version length payload{ required_buffer_size required_buffer_Manage } }		16 8 16 32 8	 uimsbf uimsbf uimsbf uimsbf uimsbf

8.2.1.3 语义

message_id: 该字段为 16 位。指示 LBM 消息的消息标识符。

version: 该字段为 8 位。指示 LBM 消息的版本。

length: 该字段为 16 位。指示 LBM 消息的字节长度。

required_buffer_size: 该字段为 32 位。指示客户端为了接收该数据，需要准备的知识层数据的缓存的字节大小。

required_buffer_Manage: 该字段为 8 位。指示客户端管理知识层数据缓存的方法，该字段取值为‘0’表示使用 FIFO 方法，取值为‘1’表示使用 LFU 方法，取值为‘2’表示使用 LRU 方法等等。

8.2.2 LBM 反馈消息

8.2.2.1 定义

LBM反馈消息用于将知识层数据缓存的管理操作反馈给服务端，告知客户端侧不可用的，需重新传输知识层数据。

8.2.2.2 语法

LBM反馈消息的语法见表5。

表5 LBM 反馈消息语法

语法	值	位数	类型
LBM_feedback_message() { message_id version length payload{ unavailable_mfu_number for(i=0;i<N;i++){ asset_id() sample_id mfu_id } } }	N	16 8 16 32 32 32	uimsbf uimsbf uimsbf uimsbf uimsbf uimsbf

8.2.2.3 语义

message_id: 该字段为 16 位。指示 LBM 反馈消息的消息标识符。

version: 该字段为 8 位。指示 LBM 反馈消息的版本。

length: 该字段为 16 位。指示 LBM 反馈消息的字节长度。

unavailable_mfu_number: 该字段为 32 位。指示知识层数据缓存中不可用的数据所属的 MFU 的数目。

asset_id(): 该字段指示第 i 个不可用 MFU 所属的资源标识符编号，该字段位数及具体用法见 T/AI 114-2021。

sample_id: 该字段为 32 位。指示第 i 个不可用 MFU 所属的样本编号。

mfu_id: 该字段为 32 位。指示第 i 个不可用 MFU 的编号。

8.3 包含非对齐 MFU 依赖关系的样本格式

8.3.1 定义

包含非对齐的MFU依赖关系的样本格式有效地支持了序列图像和知识图像依赖关系的描述及对知识图像进行高效的管理。

8.3.2 语法

```
aligned(8) class SMTHRefSample extend SMTHSample {  
    referenceMFUInfo();  
}  
  
aligned(8) class referenceMFUInfo extends Box('refm') {  
    bit(1) reference_Sample_flag;  
    bit(1) is_firstMFUinSample_flag;  
    bit(6) reserved0;  
    if (reference_Sample_flag&&is_firstMFUinSample_flag) {  
        unsigned int(32) depended_Sample_id;  
    }  
}
```

8.3.3 语义

reference_sample_flag: 指示当前 MFU 所属的 sample 是否依赖 MFU, 值 ‘0’ 意味着不参考。

is_firstMFUinSample_flag: 指示当前 MFU 是否为所属 sample 的第一个 MFU, 值 ‘0’ 表示不是。

depended_sample_id: 指示参考的 sample 的编号。

8.4 Asset 组描述符

8.4.1 定义

SMT 中定义了 Asset 组描述符, 用于指示同一个 SMT Package 中 Asset 的关联关系。在 SMT 中, 原来只有四种关系类型, 分别是依赖关系、组合关系、等同关系和相似关系, 相应的 flag 分别是 dependency_flag、composition_flag、equivalence_flag 和 similarity_flag。在 SMT 中添加的新的关系类型是非对齐时间段的知识位流依赖关系类型, 对应的 flag 是 library_flag, 该关系类型是用来描述当前 Asset 与非对齐时间段的知识位流 Asset 的依赖关系。

8.4.2 语法

Asset 组描述符的语法见表6。

表6 Asset 组描述符语法

语法	值	位数	类型
Asset_group_descriptor() { descriptor_tag descriptor_length reserved dependency_flag composition_flag equivalence_flag similarity_flag library_flag if(dependency_flag) { num_dependencies for(i = 0; i < N1; i++) { asset_id() } } if(composition_flag) { num_compositions for(i = 0; i < N2; i++) { asset_id() } } if(equivalence_flag) { equivalence_selection_level num_equivalences for(i = 0; i < N3; i++) { asset_id() equivalence_selection_level } } }	‘1111’ N1 N2 N3	16 16 4 1 1 1 1 1 1 8 8 8 8 8	uimsbf uimsbf blsbf blsbf blsbf blsbf blsbf blsbf uimsbf uimsbf uimsbf uimsbf uimsbf

表6 (续)

语法	值	位数	类型
<pre> } if(similarity_flag) { similarity_selection_level num_similarities for(i = 0; i < N4; i++) { asset_id() similarity_selection_level } } if(library_flag) { num_libraries for(i = 0; i < N5; i++) { asset_id() } } } </pre>	N4	8 8	uimsbf uimsbf
	N5	8	uimsbf

8.4.3 语义

descriptor_tag: 该字段为 16 位。用于指示此类型描述符的标签值。

descriptor_length: 该字段为 16 位。指示此描述符的字节长度，从下一个字段计算至最后一个字段。

dependency_flag: 该字段为 1 位。指示在此描述符中是否需要添加依赖关系。值‘0’意味着不需要添加。

composition_flag: 该字段为 1 位。指示在此描述符中是否需要添加组合关系。值‘0’意味着不需要添加。

equivalence_flag: 该字段为 1 位。指示在此描述符中是否需要添加等同关系。值‘0’意味着不需要添加。

similarity_flag: 该字段为 1 位。指示在此描述符中是否需要添加相似关系。值‘0’意味着不需要添加。

library_flag: 该字段为 1 位。指示在此描述符中是否需要添加非对齐时间段的知识位流依赖关系。值‘0’意味着不需要添加。

num_dependencies: 该字段为 8 位。指示此描述符所描述的 Asset 所依赖的 Asset 的数目。

num_compositions: 该字段为 8 位。指示与此描述符所描述的 Asset 有组合关系的 Asset 的数目。

equivalence_selection_level: 该字段为 8 位。指示所对应的 Asset 在等同关系组中的呈现等级。‘0’值表示该 Asset 被默认呈现。当默认 Asset 无法被选择时，拥有呈现等级较小的 Asset 会作为替代被选择和呈现。

num_equivalences: 该字段为 8 位。指示与此描述符所描述的 Asset 有等同关系的 Asset 的数目。

similarity_selection_level: 该字段为 8 位。指示所对应的 Asset 在相似关系组中的呈现等级。‘0’值表示该 Asset 被默认呈现。当默认 Asset 无法被选择时，拥有呈现等级较小的 Asset 会作为替代被选择和呈现。

num_similarities: 该字段为 8 位。指示与此描述符所描述的 Asset 有相似关系的 Asset 的数目。

num_libraries: 该字段为 8 位。指示此描述符所描述的 Asset 所依赖的非对齐时间段的知识位流 Asset 的数目。

asset_id: 该字段指示 Asset 的标识符，表 6 所定义的 Asset 组描述符中的 **asset_id**：

- a) 当 Asset 组描述符用于指示依赖关系时，**asset_id** 字段指示此描述符所描述的 Asset 所依赖的 Asset 的标识符，同时，此描述符中提供的 Asset 标识符顺序与其内部编码依赖层次相对应；
- b) 当 Asset 组描述符用于指示组合关系时，**asset_id** 字段指示与此描述符所描述的 Asset 有组合关系的 Asset 的标识符；
- c) 当 Asset 组描述符用于指示等同关系时，**asset_id** 字段指示与此描述符所描述的 Asset 有等同关系的 Asset 的标识符；
- d) 当 Asset 组描述符用于指示相似关系时，**asset_id** 字段指示与此描述符所描述的 Asset 有相似关系的 Asset 的标识符；
- e) 当 Asset 组描述符用于指示非对齐时间段的知识位流依赖关系时，**asset_id** 字段指示与此描述符所描述的 Asset 有非对齐时间段的知识位流依赖关系的 Asset 的标识符。

该字段位数及具体用法见 T/AI 114-2021。

8.5 彩色信息描述符

8.5.1 定义

SMT传输的AVS3视频编码位流的彩色信息描述符用于描述AVS3视频的彩色信息。

8.5.2 语法

AVS3彩色信息描述符的语法见表7。

表7 AVS3 彩色信息描述符语法

语法	值	位数	类型
Colour_info_descriptor () { descriptor_tag descriptor_length num_colour_asset for(i=0; i<N; i++){ asset_id() colour primaries transfer characteristics matrix_coefficients } }	N	16 16 8 8 8 8	uimsbf uimsbf uimsbf uimsbf uimsbf uimsbf

8.5.3 语义

descriptor_tag: 该字段为 16 位。标识符，用于标识描述符的类型，其取值为 0x0009 时指示描述符为彩色信息描述符。

descriptor_length: 该字段为 16 位。标识符的长度，单位为字节。指示 AVS3 彩色信息描述符的字节长度。

num_colour_asset: 该字段为 8 位。指示具有彩色信息描述的媒体资源数量，所述彩色信息描述包括彩色三基色、光电转移特性和彩色信号转换矩阵。

asset_id: 该字段指示具有彩色信息描述的媒体资源的标识符,该字段位数及具体用法参见 T/AI 114-2021。

colour_primaries: 该字段为 8 位。指示彩色三基色,说明源图像三基色的色度坐标,其取值应与 T/AI 109.2-2021 中的 colour_primaries 的值相同。

transfer_characteristics: 该字段为 8 位。指示源图像的光电转移特性,其取值应与 T/AI 109.2-2021 中的 transfer_characteristics 的值相同。

matrix_coefficients: 该字段为 8 位。指示彩色信号转换矩阵,说明从红绿蓝三基色转化为亮度和色度信号时所采用的转换矩阵,其取值应与 T/AI 109.2-2021 中的 matrix_coefficients 的值相同。

9 传输流和节目流技术要求

9.1 基本要求

AVS3 视频编码位流支持传输流和节目流的封装格式,编码结构与参数应符合第 9 章定义。

AVS3 视频基本流的传输流和节目流封装和解封装见附录 C。

AVS3 视频流应符合以下约束:

- AVS3 视频流应是 ISO/IEC 13818-1:2022 中节目的一个节目元素,基本流的 stream_type 字段值应等于 ‘0xD4’;
- T/AI 109.2-2021 中定义的用于解码 AVS3 视频流的序列头应存在于承载该视频流的基本流中;
- 在 stream_type 等于 ‘0xD4’ 且具有 ISO/IEC 13818-1:2022 中层次描述符的基本流中,层次描述符中的 hierarchy_type 字段值应等于 ‘3’ (时域可扩展性);
- AVS3 视频流的最高“级别”以及所符合的“档次”等编码参数应使用 AVS3 视频流描述符标识。如果存在与 AVS3 视频流相关联的 AVS3 视频流描述符,则该描述符应包含在节目映射表中相应基本流条目的描述符循环中。

包含主位流及其依赖的知识位流的 AVS3 视频流应符合以下约束:

- 当 AVS3 视频流被封装为传输流时,主位流及其依赖的知识位流应被封装在同一个传输流中;
- 当 AVS3 视频流被封装为节目流时,主位流及其依赖的知识位流应被封装在同一个节目流中。

9.2 PES 分组

9.2.1 流标识

AVS3 视频流应作为 PES_packet_data_bytes 携带在 PES 分组数据包中,并通过节目映射表中分配的 stream_type 字段值 (0xD4) 标识。

以下约束适用于 AVS3 视频的 PES 数据包使用的 stream_id:

- a) PES 分组中的 stream_id 取值 ‘1111 1101’ (extended_stream_id), 表示 PES 分组数据包采用扩展语法标识 AVS3 视频流;
- b) PES 分组包头中 stream_id_extension_flag 取值 ‘0’, stream_id_extension 的取值范围应从 ‘0x41’ 到 ‘0x4F’, stream_id_extension 字段的取值设置为 ‘0x41’ 用于表示 AVS3 视频主位流,该字段取值设置为 ‘0x42’ 用于表示 AVS3 视频知识位流。

注: ‘0x41’ 到 ‘0x4F’ 范围内的其他值可能后续被 AVS 标准使用。

9.2.2 数据对齐指示

AVS3 视频流的 PES 分组使用 `data_alignment_indicator` 和 9.3.4 中定义的数据流对齐描述符 `data_stream_alignment_descriptor` 来标识 PES 分组包头后数据的对齐方式：

- 如果 `data_alignment_indicator` 值为 ‘1’，并且存在 9.3.4 中定义的数据流对齐描述符 `data_stream_alignment_descriptor`，表示 PES 分组包头之后紧跟着该数据流对齐描述符所指出的视频起始码；
- 如果 `data_alignment_indicator` 值为 ‘1’ 且不存在 9.3.4 中定义的数据流对齐描述符 `data_stream_alignment_descriptor`，则使用表 11 中 `alignment_type` 取值 ‘01’ 时所指示的对齐方式；
- 当 `data_alignment_indicator` 值为 ‘0’ 时，表示没有定义是否有任何对齐方式。

9.2.3 解码时间约束

对于知识位流，PES 分组包的包头中的 `PTS_DTS_flags` 字段设置为 ‘11’，此时知识位流 PES 分组包的包头中携带的 PTS 的值应被忽略，解码后的知识图像应不被用于显示。

知识位流的 PES 分组包的包头中携带的 DTS 应小于或等于 AVS3 视频的主位流的依赖该 PES 分组包的所有 PES 分组包中携带的最小的 DTS。

9.3 节目和节目元素描述符

9.3.1 节目和节目元素描述符中各字段的语义定义

以下语义适用于从 9.3.2 到 9.3.5 中定义的描述符。

描述符标签字段 `descriptor_tag`

该字段为 8 位，用于标识每一描述符。

数据流对齐描述符 `data_stream_alignment_descriptor` 和 AVS3 视频流描述符 `AVS3_video_descriptor` 的标签值见表 8。TS 或 PS 栏中 ‘X’ 表示该描述符可分别用于传输流或节目流。注意，描述符字段含义可能取决于它用于的流。以下描述符语义对每种情况作了规定。

表8 节目和节目元素描述符

描述符标签值	TS	PS	标识
6	X	X	<code>data_stream_alignment_descriptor</code>
209	X	X	<code>AVS3_video_descriptor</code>

描述符长度字段 `descriptor_length`

该字段为 8 位。规定了紧跟在 `descriptor_length` 字段之后的描述符的字节数。

9.3.2 AVS3 视频流描述符

AVS3 视频流描述符给出了标识 T/AI 109.2-2021 中定义的 AVS3 视频流的编码参数（例如“档次”和“级别”）的基本信息，见表 9。对于 AVS3 视频时域分层子位流，AVS3 视频流描述符提供在其所适用的基本流中的 AVS3 视频最高时域层级的信息。

表9 AVS3 视频流描述符

语 法	位 数	助 记 符
AVS3_video_descriptor () {		
descriptor_tag	8	uimsbf
descriptor_length	8	uimsbf
profile_id	8	uimsbf
level_id	8	uimsbf
multiple_frame_rate_flag	1	bslbf
frame_rate_code	4	uimsbf
sample_precision	3	uimsbf
chroma_format	2	uimsbf
temporal_id_flag	1	bslbf
td_mode_flag	1	bslbf
library_stream_flag	1	uimsbf
library_picture_enable_flag	1	uimsbf
reserved	2	bslbf
transfer_characteristics	8	uimsbf
matrix_coefficients	8	uimsbf
reserved	8	bslbf
}		

9.3.3 AVS3 视频流描述符中各字段的语义定义

档次标识 **profile_id**

该字段为 8 位。表示视频位流的档次，该字段应根据 T/AI 109.2-2021 中定义的 **profile_id** 进行编码。

级别标识 **level_id**

该字段为 8 位。表示视频位流的级别。该字段应根据 T/AI 109.2-2021 中定义的 **level_id** 进行编码。

复合帧速率标志 **multiple_frame_rate_flag**

该字段为 1 位，置 ‘1’ 时表示视频流中可能有多个帧速率，置 ‘0’ 时表示只有单一帧速率。

帧速率码 **frame_rate_code**

该字段为 4 位，该字段应根据 T/AI 109.2-2021 中定义的 **frame_rate_code** 进行编码。
multiple_frame_rate_flag 字段置为 ‘1’ 时，特定的帧速率指示还允许某些其它帧速率出现在视频流中，见表 10。

表10 帧速率码指示的帧速率

帧速率	同时允许的帧速率
23.976	
24.0	23.976
25.0	
29.97	23.976
30.0	23.976 24.0 29.97
50.0	25.0
59.94	23.976 29.97
60.0	23.976 24.0 29.97 30.0 59.94
100.0	50.0
119.88	59.94
120.0	59.94 60.0 119.88

样本精度字段 sample_precision

该字段为 3 位。规定亮度和色度样本的精度。该字段应根据 T/AI 109.2-2021 中定义的 sample_precision 进行编码。

色度格式字段 chroma_format

该字段为 2 位。规定色度分量的格式。该字段应根据 T/AI 109.2-2021 中定义的 chroma_format 进行编码。

时间层标识允许标志 temporal_id_flag

该字段为 1 位。表示视频流是否允许使用时间层标识。该字段应根据 T/AI 109.2-2021 中定义的 temporal_id_enable_flag 进行编码。

立体视频模式标志 td_mode_flag

该字段为 2 位。表示视频流是单目视频流或多视点视频流。该字段应根据 T/AI 109.2-2021 中定义的 td_mode_flag 进行编码。

知识位流标志 library_stream_flag

该字段为 1 位。指示在包含此描述符的节目特定信息的生效范围内的基本流是否为知识位流。值为‘1’表示该基本流是知识位流；值为‘0’表示该基本流是主位流。

知识图像允许标志 library_picture_enable_flag

该字段为 1 位。指示描述符对应的主位流中是否存在帧间预测图像使用知识图像作为参考图像。值为‘1’表示主位流中的帧间预测图像使用知识图像作为参考图像；值为‘0’表示主位流中的帧间预测图像不使用知识图像作为参考图像。

彩色三基色 colour_primaries

该字段为 8 位。说明视频流中源图像三基色的色度坐标。该字段应根据 T/AI 109.2-2021 中定义的 colour_primaries 进行编码。

光电转移特性 transfer_characteristics

该字段为 8 位。说明视频流中源图像的光电转移特性。该字段应根据 T/AI 109.2-2021 中定义的 transfer_characteristics 进行编码。

彩色信号转换矩阵 `matrix_coefficients`

该字段为 8 位。说明从红绿蓝三基色转换为亮度和色度信号时采用的转换矩阵。该字段应根据 T/AI 109.2-2021 中定义的 `matrix_coefficients` 进行编码。

9.3.4 数据流对齐描述符

ISO/IEC 13818-1:2022中规定的的数据流对齐描述符描述了AVS3视频基本流中存在的对齐类型。如果 PES 分组包头中的`data_alignment_indicator`的值为‘1’并且存在数据流对齐描述符，则需要由描述符指示对齐类型，见表11。

表11 AVS3 视频流对齐类型

对齐类型	描述
00	保留
01	AVS3 视频存取单元
02	AVS3 视频片
03	AVS3 视频存取单元或 AVS3 视频片
04	AVS3 视频序列
05 - 255	保留

9.3.5 数据流对齐描述符中各字段的语义定义

对齐类型字段 `alignment_type`

该字段是数据流对齐描述符中的 8 位字段，它描述了 AVS3 视频基本流中存在的对齐类型。

当 PES 分组包头中的 `data_alignment_indicator` 为‘1’时 T/AI 109.2-2021 视频的对齐类型见表 11。如果 PES 分组包是一个视频流的起始，对齐应发生在第一个序列头的起始码处。

一个AVS3视频存取单元包含一个图像的所有编码数据，以及紧随其后的任意填充数据，直到下一个存取单元的起始位置为止且不包含下一个存取单元。如果一个图像前面没有视频序列起始码（`video_sequence_start_code`），那么这个图像的存取单元以图像起始码开始。如果一个图像前面有视频序列起始码，那么这个图像的存取单元以这些起始码中的第一个起始码的第一个字节开始。如果一个图像是码流中在视频序列结束码（`video_sequence_end_code`）之前的最后一个图像，那该编码图像的最后一个字节和视频序列结束码之间的所有字节（包括视频序列结束码）属于该编码图像的存取单元。

9.4 DPB 管理

使用本文件携带 AVS3 视频编码位流不影响 DPB 的大小。当在 STD 中解码 AVS3 视频编码位流时，DPB 的大小由 T/AI 109.2-2021 定义。DPB 应按照 T/AI 109.2-2021 中 9.2.4 进行管理。在一个 AVS3 视频存取单元被解码之后，即在该 AVS3 视频存取单元从 CPB 中被移除的时刻，该被解码的 AVS3 视频存取单元立即进入 DPB。

一个被解码的 AVS3 主位流的存取单元在 DPB 输出时间所指示的时刻被显示，一个被解码的 AVS3 知识位流的存取单元在其被解码之后立即从 DPB 中输出。DPB 输出时间是执行 T/AI 109.2-2021 中 9.2.6 中的图像输出操作的时间所指示的时刻。

如果 AVS3 视频位流不能提供足够信息用于决定 AVS3 主位流的存取单元的 CPB 移除时间和 DPB

输出时间，那么这些时间戳应在 STD 模型中根据 PTS 和 DTS 时间戳决定，方法如下：

- a) AVS3 主位流的存取单元 n 的 CPB 移除时间是由 $DTS(n)$ 指示的时刻，其中 $DTS(n)$ 是 AVS3 主位流的存取单元 n 的 DTS 值；
- b) AVS3 主位流的存取单元 n 的 DPB 输出时间是由 $PTS(n)$ 指示的时刻，其中 $PTS(n)$ 是 AVS3 主位流的存取单元 n 的 PTS 值。

如果 AVS3 视频位流不能提供足够信息用于决定 AVS3 知识位流的存取单元的 CPB 移除时间，那么该时间戳应在 STD 模型中根据 DTS 时间戳被决定，方法为：AVS3 知识位流的存取单元 n 的 CPB 移除时间是由 $DTS(n)$ 指示的时刻，其中 $DTS(n)$ 是 AVS3 知识位流存取单元 n 的 DTS 值。

9.5 面向 AVS3 的 T-STD 扩展

对于一个符合本文件的节目，当该节目中的位流是 AVS3 视频编码位流或包含最低时域层级的 AVS3 视频编码位流时，应符合 T-STD 模型，见图1。

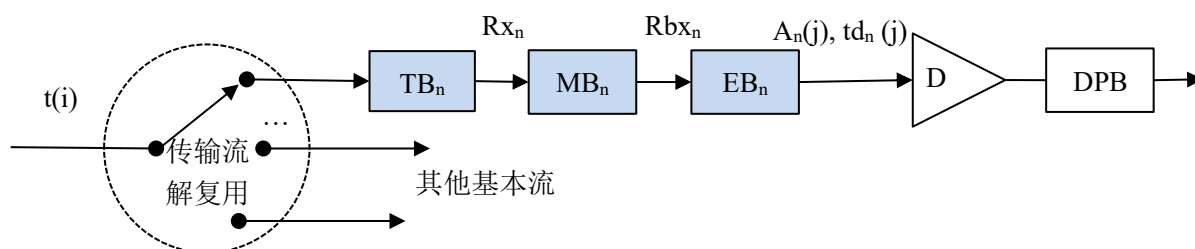


图1 面向 AVS3 的 T-STD 模型扩展

图中符号说明：

$t(i)$ ——表示传输流的第 i 个字节进入系统目标解码器的时间，以秒为单位。

TB_n ——基本流 n 的传输缓存。

TBS ——传输缓存 TB_n 的大小，以字节为单位。

MB_n ——基本流 n 的复用缓存。

MBS_n ——复用缓存 MB_n 的大小，以字节为单位。

EB_n ——AVS3 视频编码位流的基本流缓存。

j ——AVS3 视频编码位流中 AVS3 视频存取单元的索引。

$A_n(j)$ ——AVS3 视频编码位流的第 j 个存取单元。

$td_n(j)$ —— $A_n(j)$ 在系统目标解码器中的解码时间，以秒为单位。

R_{xn} ——从传输缓存 TB_n 到复用缓存 MB_n 的传输速率。

R_{bxn} ——从复用缓存 MB_n 到基本流缓存 EB_n 的传输速率。

注：图中阴影部分是 T-STD 模型包括的缓存，这些缓存位于解码器之前。

9.5.1 缓存管理

缓存管理符合下列约束：

- a) 仅有一个传输缓存 TB_n 用于接收 AVS3 视频编码位流或包含最低时域层级的 AVS3 视频编码位流，其大小 TBS 固定为 512 字节；

- b) 仅有一个复用缓存 MB_n 用于接收 AVS3 视频编码位流或包含最低时域层级的 AVS3 视频编码位流, 该复用缓存的大小 MBS_n 约束见公式 (1):

$$MBS_n = BS_{mux} + BS_{oh} + MaxCPB[level_id] - BBS \quad \dots\dots\dots (1)$$

式中:

MBS_n ——复用缓存, 单位为位;

BS_{oh} ——包头缓存, 单位为位, 定义见公式 (2):

$$BS_{oh} = (1/750) \times \max\{MaxBR[level_id], 2\,000\,000\} \quad \dots\dots\dots (2)$$

式中:

BS_{mux} ——额外复用缓存, 单位为位, 定义见公式 (3):

$$BS_{mux} = 0.004 \times \max\{MaxBR[level_id], 2\,000\,000\} \quad \dots\dots\dots (3)$$

式中:

$MaxCPB[level_id]$ ——最大 BBV 缓冲区大小, 单位为位, 根据 AVS3 视频编码位流或包含最低时域层级的 AVS3 视频编码位流的级别 $level_id$ 由 T/AI 109.2-2021 中的附录 B.3 的参数限制中得到;

$MaxBR[level_id]$ ——最大位速率, 单位为位每秒, 根据 AVS3 视频编码位流或包含最低时域层级的 AVS3 视频编码位流的级别 $level_id$ 由 T/AI 109.2-2021 中的附录 B.3 的参数限制中得到;

BBS ——BBV 缓冲区大小, 单位为位, 根据 AVS3 视频编码位流或包含最低时域层级的 AVS3 视频编码位流中由 T/AI 109.2-2021 中 7.1.2.2 和 7.2.2.2 定义的序列头参数 bbv_buffer_size 得到。

- c) 仅有一个基本流缓存 EB_n 用于接收分层描述子对应的基本流集合中的所有基本流, 该基本流缓存的总大小 EBS_n 定义见公式 (4):

$$EBS_n = BBS \quad \dots\dots\dots (4)$$

式中:

EBS_n ——基本流缓存的总大小, 单位为位;

BBS ——BBV 缓冲区大小, 单位为位, 根据 AVS3 视频编码位流或包含最低时域层级的 AVS3 视频编码位流中由 T/AI 109.2-2021 中 7.1.2.2 和 7.2.2.2 定义的序列头参数 bbv_buffer_size 得到。

- d) 从 TB_n 到 MB_n 的传输应按照如下方式执行:

当 TB_n 中没有数据时, R_{x_n} 设为 0; 否则, 计算方式见公式 (5):

$$R_{x_n} = BitRate \times 400 \quad \dots\dots\dots (5)$$

式中:

R_{x_n} ——从传输缓存 TB_n 到复用缓存 MB_n 的传输速率, 单位为位每秒;

$BitRate$ ——AVS3 视频编码位流或包含最低时域层级的 AVS3 视频编码位流中的位率参数, 单位为位每秒。当 AVS3 视频编码位流不支持时间分层时, 位率参数根据 T/AI 109.2-2021 中 7.1.2.2 和 7.2.2.2 定义的位率低位和位率高位计算得到; 当 AVS3 视频编码位流支持时间分层时, 位率参数根据 T/AI 109.2-2021 中 7.1.2.7 和 7.2.2.7 定义的时间层位率低位和时间层位率高位计算得到。

- e) 从 MB_n 到 EB_n 的传输应按照公式 (6) 执行:

$$R_{bx_n} = MaxBR[level_id] \quad \dots\dots\dots (6)$$

式中：

R_{bx_n} ——从复用缓存 MB_n 到基本流缓存 EB_n 的传输速率，单位为位每秒；

$MaxBR[level_id]$ ——最大位速率 $MaxBR[level_id]$ ，单位为位每秒，根据 AVS3 视频编码位流或包含最低时域层级的 AVS3 视频编码位流的级别 $level_id$ 由 T/AI 109.2-2021 中的附录 B.3 的参数限制中得到。

如果 MB_n 中存在 PES 分组包负载数据且 EB_n 没有充满，PES 分组包负载以速率 R_{bx_n} 从 MB_n 被传输到 EB_n 。如果 EB_n 已经被充满， MB_n 中的数据不应被移除。当一个字节的数据从 MB_n 被传输到 EB_n 时， MB_n 中位于该字节之前的 PES 分组包头的所有字节被立即移除和丢弃。当 MB_n 中不存在 PES 分组包负载数据时，没有数据被从 MB_n 中移除。所有进入 MB_n 的数据都应从其中离开。所有离开 MB_n 的 PES 分组包负载数据都要立即进入 EB_n 。

9.5.2 STD 延时

STD 延时是数据通过 STD 中的 TB_n 、 MB_n 和 EB_n 缓存的时间，符合 T/AI 109.2-2021 的任何数据的 STD 延时应符合约束：对于所有的 j 对应的存取单元 $A_n(j)$ 中的所有字节 i ， $td_n(j) - t(i) \leq 10$ (s)。

9.5.3 缓存管理条件

传输流的构建应符合如下的缓存管理条件：

- 每个 TB_n 不应上溢，且应在每一秒中至少有一次处于被清空的状态；
- 每个 MB_n 、 EB_n 和 DPB 不应上溢；
- EB_n 不应下溢。当 AVS3 视频存取单元 $A_n(j)$ 中一个或多个字节在解码时间 $td_n(j)$ 不存在于 EB_n 中时， EB_n 对于 $A_n(j)$ 发生下溢。

附 录 A
(规范性)
MIME 类型的 ‘codecs’ 参数

A.1 基本要求

当使用IETF RFC 6381定义的MIME类型的 ‘codecs’ 参数时，如果MIME类型对应本文件中定义的文件格式且 ‘codecs’ 参数值起始于5.3.1和5.3.2的样本入口类型，则 ‘codecs’ 参数的子参数应符合A.2定义的规定。

A.2 AVS3 视频编码参数

如果 ‘codecs’ 参数第一个元素的取值代码对应于T/AI 109.2-2021定义的编解码，即5.3.1和5.3.2中定义的AVS3视频编码（‘avs3’或‘lav3’），且对应的轨道表示AVS3视频编码位流，则之后元素分别取值为AVS3解码器配置记录中定义的“档次”和“级别”字段值。‘codecs’ 参数的不同元素之间以句号隔开（“.”），在所有数字编码中，可以省略前导零。

- a) 档次（profile_id）： ‘codecs’ 参数的第二个子参数，编码为十六进制数；
- b) 级别（level_id）： ‘codecs’ 参数的第三个子参数，编码为十六进制数。

AVS3 视频编码参数示例如下：

示例：

codecs= avs3.20.54

——表示以基准 8 位档次（Main 8bit profile），级别 8.0.60 编码的 AVS3 视频编码位流。

codecs= avs3.22.68

——表示以基准 10 位档次（Main 10bit profile），级别 10.0.120 编码的 AVS3 视频编码位流。

codecs= avs3.30.54

——表示以加强 8 位档次（High 8bit profile），级别 8.0.60，编码的 AVS3 视频编码位流。

codecs= avs3.32.68

——表示以加强 10 位档次（High 10bit profile），级别 10.0.120，编码的 AVS3 视频编码位流。

附录 B (资料性) 知识位流编码媒体数据传输

B.1 SMT智能媒体传输

一种将媒体数据分段并传输的方法见图 B.1，将包含主位流和知识位流的媒体数据封装在一个包裹中，该包裹包含资源 1 和资源 2 两个资源，同时还包含一个组成信息（Composition Information, CI）。每个资源包含一个 CEU，每个 CEU 包含主位流或知识位流的一种数据，例如资源 1 的 CEU1 包含主位流，资源 2 的 CEU2 包含知识位流。组成信息记录了资源之间的时域、空域或依赖关系等信息，见图 B.1，组成信息描述了资源 1 对资源 2 的依赖性。每个 CEU 中包含了至少一个 MFU，并由提示轨道（hint track）描述 MFU 在 CEU 中的分段信息，见图 B.1，CEU2 被分段为 MFU1 和 MFU4，而 CEU1 被分段为 MFU2、MFU3、MFU5-7，其中虚线表示 MFU 之间的依赖关系，资源 1 中 MFU2、MFU3 和 MFU6 依赖资源 2 中 MFU1，资源 1 中 MFU5 和 MFU7 依赖资源 2 中 MFU4。相互依赖的 MFU 需要被同步传输到客户端，见图 B.1，实线箭头在时间线上描述 MFU 的传输时间。

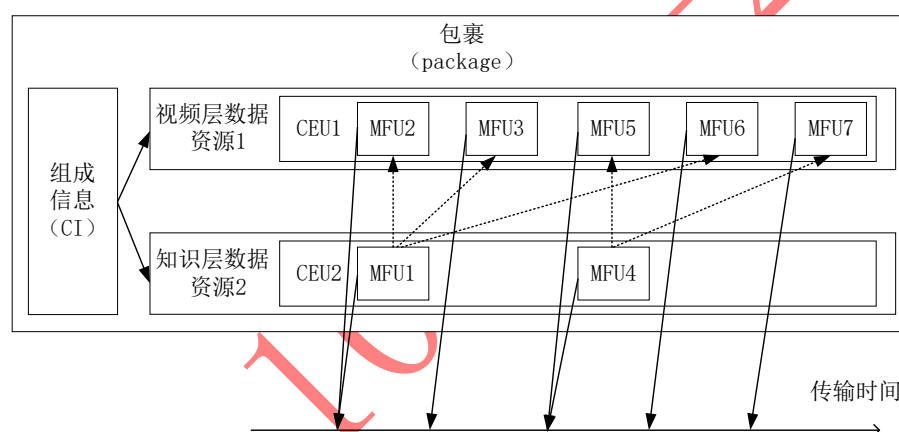


图 B.1 媒体数据分段传输方法

在确定 MFU 的依赖关系并能够唯一定位 MFU 之后，当传输 MFU 时，需要同步传输有依赖关系的 MFU。MFU 传输流程见图 B.2，首先根据当前传输顺序从资产 1 的主位流中获取当前 MFU，例如图 B.1 中资产 1 中的 MFU2。根据当前 MFU 的样本信息，判断当前 MFU 是否依赖一个 Dependent MFU (DMFU)。如果当前 MFU 不依赖 DMFU，那么传输当前 MFU 并继续按照顺序获取下一个 MFU 或终止传输；如果当前 MFU 依赖 DMFU，那么根据当前 MFU 中描述的 DMFU 的索引，从资产 2 中的数据中获取 DMFU。由于多个 MFU 可能依赖同一个 MFU，为了避免 DMFU 的重复传输，在传输 DMFU 时，需要考虑三种情况，以判断 DMFU 在客户端的可用性，见图 B.2。在一种情况下，根据 DMFU 的历史传输列表，当前 MFU 依赖的 DMFU 没有被传输过，那么需要将 DMFU 和当前 MFU 同步传输，见图 B.1，MFU1 和 MFU2 需要被同步传输；在另一种情况下，根据 DMFU 的历史传输列表，当前 MFU 依赖的 DMFU 已经被传输过，那么只需要传输当前 MFU 而不需要传输 DMFU，见图 B.1，MFU3、MFU6 依赖的 MFU1 已经与 MFU2 同步传输，MFU7 依赖的 MFU4 已经与 MFU5 同步传输；在又一种情况下，根据 DMFU 的历史传输列表，当前 MFU 依赖的 DMFU 已经被传输过，但是，根据客户端反馈的信令消息，该 DMFU 由于使用频次、存储、管理方法等多种可能的原因，在客户端已经不可用，此时需要

将 DMFU 和当前 MFU 同步传输，例如客户端只能缓存 1 个 MFU，当传输 MFU5 时，同步传输的 MFU4 替换掉已有的 MFU1，这导致 MFU1 的不可用，因此在传输 MFU6 时，需要同步再次传输 MFU1。

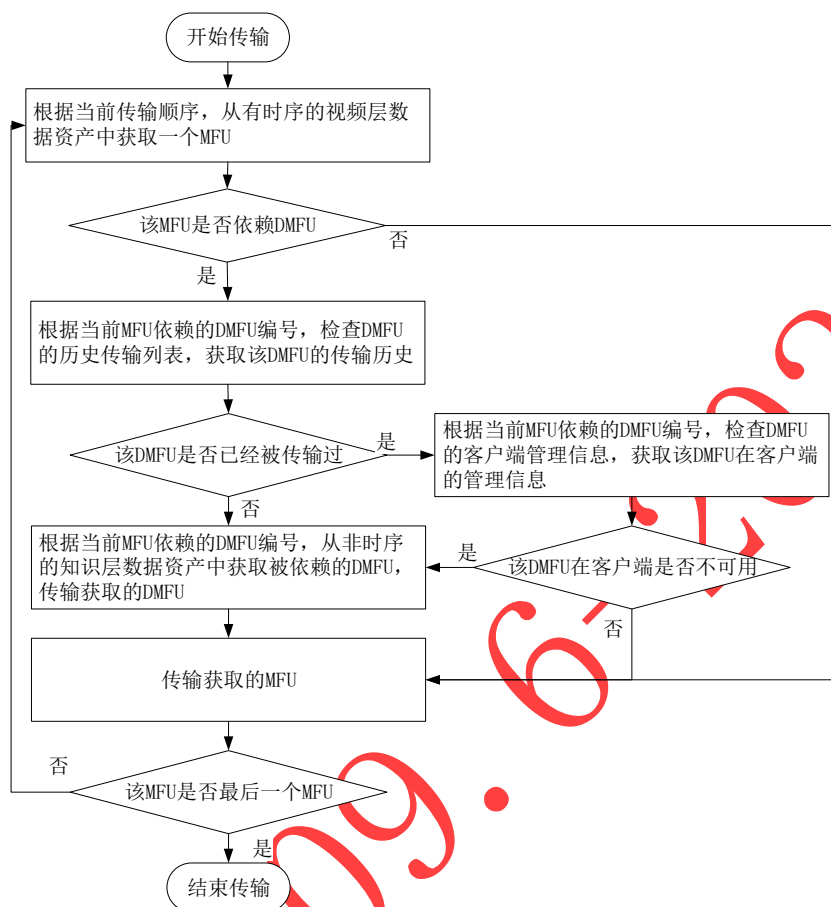


图 B.2 MFU 传输流程

B.2 MFU传输

一种传输包含主位流和知识位流的媒体数据的方法见图 B.3，该方法用于该媒体数据中样本的重复传输。

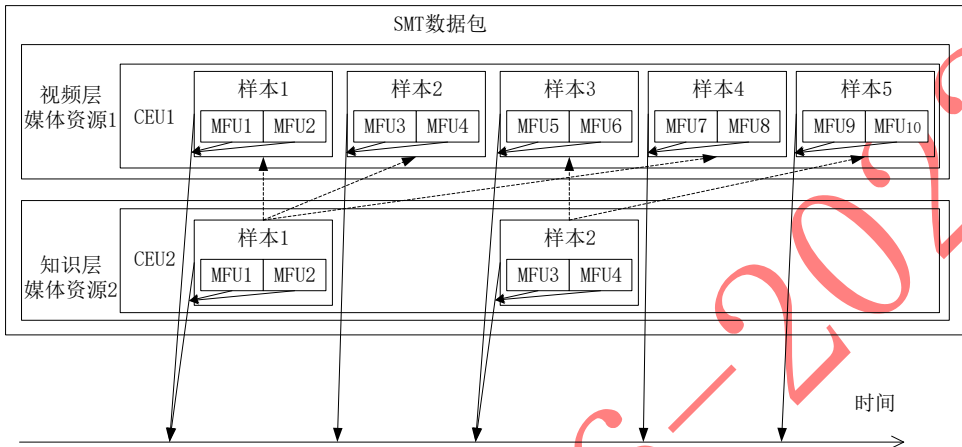


图 B.3 媒体数据样本重复传输方法

包含主位流和知识位流的媒体数据结构关系的样本见图 B.3，该媒体数据被封装在一个 SMT 数据包内，数据包包括：

- a) 一个主位流媒体资源；
- b) 一个知识位流媒体资源。

由 SMT 提示轨道里面的提示样本描述两个资源间的样本依赖信息，主位流媒体资源中的样本根据时序被传输给客户端，并且由 SMT 提示轨道描述的其所依赖的知识位流媒体资源中的样本被同步传输，如图 B.3 实线箭头在时间线上描述的传输时间。

避免重传的操作流程见图 B.4，其关键是在传输一个知识位流的样本之前，要首先判断该样本是否已经被传输过，执行以下操作之一：

- a) 如果被传输过且还可用则不再重复传输；
- b) 如果被传输过但不可用则再次传输；
- c) 如果没被传输过，则跟随主位流样本同时传输。

其中通过缓存内容更新信令来判断是否需要重传。

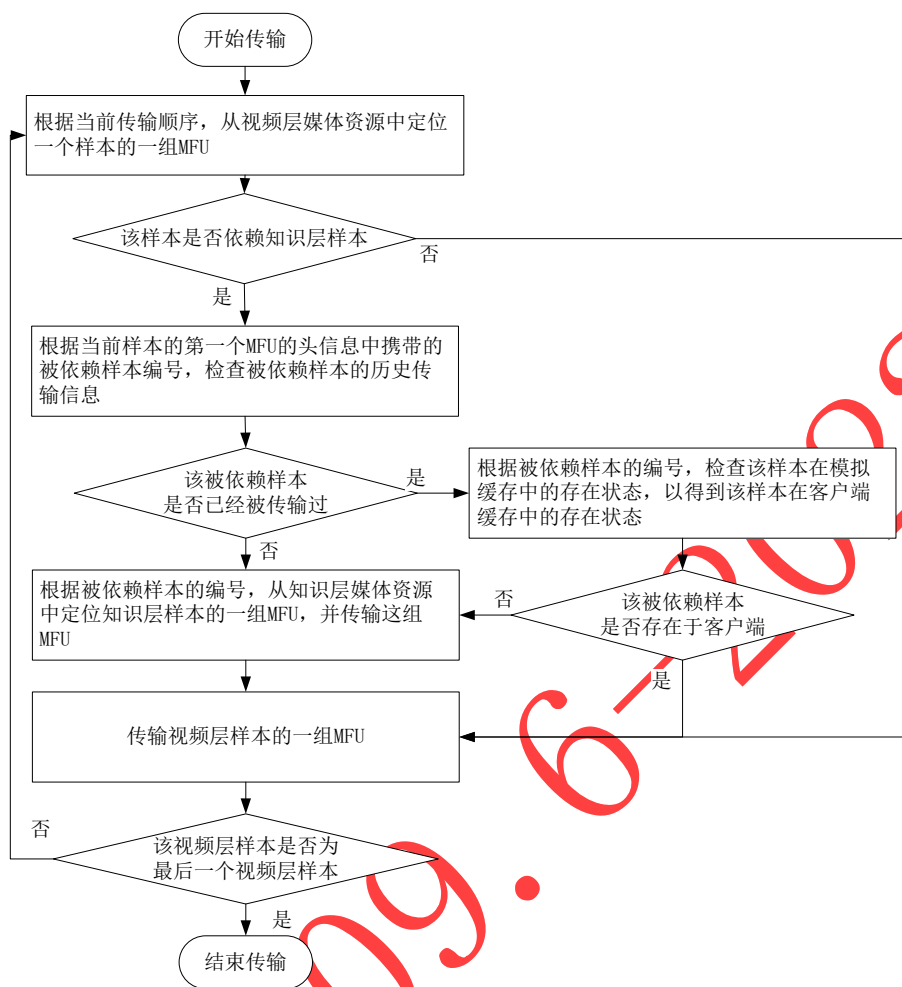


图 B.4 传输 MFU 的操作流程

附录 C

(资料性)

AVS3 视频基本流的传输流和节目流封装和解封装

C.1 传输流生成

将 AVS3 视频编码的基本流封装成传输流需要使用 AVS3 视频流描述子标识主位流以及可能存在的知识位流。

生成传输流应从视频基本流中解析语法元素用于为传输流中 PSI 表中的相关字段设定该字段应取的值。在为视频基本流所属的节目建立 PMT 时,从视频基本流的序列头中解析 `library_stream_flag` 和 `library_picture_enable_flag` 的值,方法如下:

- a) 当 `library_stream_flag` 的值为 0 且 `library_picture_enable_flag` 的值为 0 时,判断该视频基本流包含的主位流不依赖知识位流,在 PMT 中应将该主位流对应的描述子中的 `library_stream_flag` 字段的值设为 0,并将 `library_picture_enable_flag` 字段的值设为 0;
- b) 当 `library_stream_flag` 的值为 0 且 `library_picture_enable_flag` 的值为 1 时,判断该视频基本流包含的主位流依赖知识位流,在 PMT 中应将该主位流对应的描述子中的 `library_stream_flag` 字段的值设为 0,见图 C.1 操作②,并将 `library_picture_enable_flag` 字段的值设为 1;
- c) 当 `library_stream_flag` 的值为 1 时,判断该视频基本流包含知识位流,在 PMT 中应将该知识位流对应的描述子中的 `library_stream_flag` 字段的值设为 1,见图 C.1 操作③。

AVS3 视频基本流的传输流封装流程示例如下:

示例:

主位流和知识位流被认为是不同的两个基本流,主位流的 PES 分组包的包头中 `stream_id_extension` 字段的取值设置为 '0x41',知识位流的 PES 分组包的包头中 `stream_id_extension` 字段的取值设置为 '0x42'。两个基本流是一个节目中的节目元素,该节目被分配唯一的 `program_number`。承载每个基本流的传输流分组包被分别分配唯一的 PID,见图 C.1 中操作①,承载主位流的 PES 分组包的传输流分组包的 PID 值被设定为 SEQ_PID,承载知识位流的 PES 分组包的传输流分组包的 PID 值被设定为 LIB_PID。两个基本流的 PID 被分别赋值给 PMT 中各自对应的 `elementary_PID` 字段。为了获取主位流与知识位流的依赖关系,应从主位流中的参考知识图像扩展 `cross_random_access_point_reference_extension()` 或 RL 图像头的参考图像队列配置集中解析得到主位流参考的知识图像的数量和索引,并配合从知识位流的知识图像头中解析得到的知识图像索引编号,建立主位流图像和知识图像之间的依赖关系。

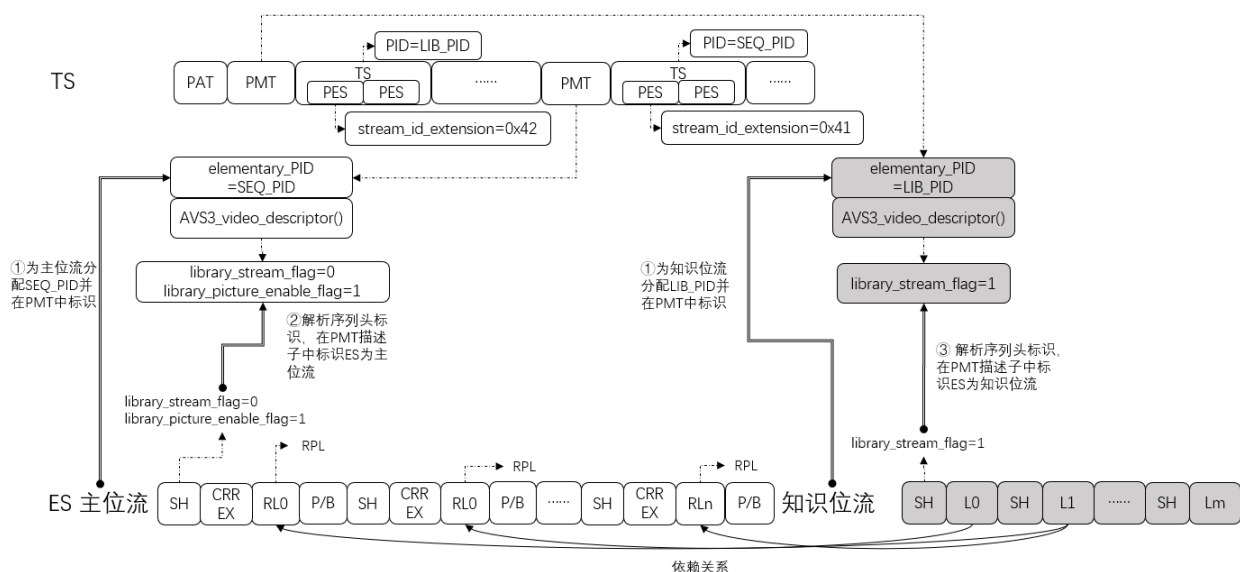


图 C.1 传输流封装流程

C.2 节目流生成

将 AVS3 视频编码的基本流封装成节目流需要使用 AVS3 视频流描述子用于标识主位流以及可能存在的知识位流。

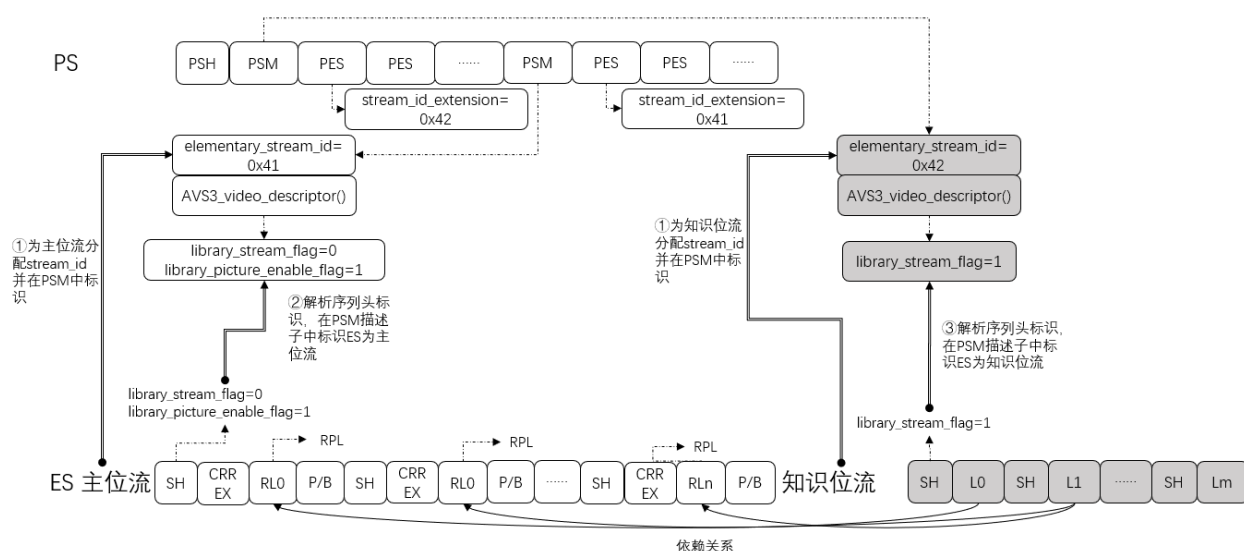
生成节目流应从视频基本流中解析语法元素用于为节目流中 PSM 中的相关字段设定该字段应取的值。从视频基本流的序列头中解析 library_stream_flag 和 library_picture_enable_flag 的值，方法如下：

- 当 library_stream_flag 的值为 0 且 library_picture_enable_flag 的值为 0 时，判断该视频基本流包含的主位流不依赖知识位流，在 PSM 中应将该主位流对应的描述子中的 library_stream_flag 字段的值设为 0，并将 library_picture_enable_flag 字段的值设为 0；
- 当 library_stream_flag 的值为 0 且 library_picture_enable_flag 的值为 1 时，判断该视频基本流包含的主位流依赖知识位流，在 PSM 中应将该主位流对应的描述子中的 library_stream_flag 字段的值设为 0，见图 C.2 操作②，并将 library_picture_enable_flag 字段的值设为 1；
- 当 library_stream_flag 的值为 1 时，判断该视频基本流包含知识位流，在 PSM 中应将该知识位流对应的描述子中的 library_stream_flag 字段的值设为 1，见图 C.2 操作③。

AVS3 视频基本流的节目流封装流程示例如下：

示例：

主位流和知识位流被认为是不同的两个基本流，主位流的 PES 分组包的包头中 stream_id_extension 字段的取值设置为 '0x41'，知识位流的 PES 分组包的包头中 stream_id_extension 字段的取值设置为 '0x42'。为了获取主位流与知识位流的依赖关系，应从主位流中的参考知识图像扩展 cross_random_access_point_reference_extension() 或 RL 图像头的参考图像队列配置集中解析得到主位流参考的知识图像的数量和索引，并配合从知识位流的知识图像头中解析得到的知识图像索引，建立主位流图像和知识图像之间的依赖关系。



注：其中虚线指向的是被包含的语法元素，实线指向的是参考实线前端的知识图像的RL图像，双实线表示根据ES的信息在PS中为相应的语法元素赋值。SH表示序列头，CRREX表示参考知识图像扩展。

图 C.2 节目流封装流程

C.3 传输流解封装

AVS3 视频基本流的传输流解封装流程如下：

a) 当传输流中只包含不依赖知识位流的主位流时，传输流的拆解操作如下：

- 1) 接收端解析 PAT 用于定位感兴趣节目的 PMT 所在的传输流分组包，解析包含 library_stream_flag 值为 0 且 library_picture_enable_flag 值为 0 的 PMT，从该 PMT 覆盖的有效范围中的传输流分组包中获取包含主位流 PES 分组包；
- 2) 接收端解析传输流分组包获得 PES 分组包，这些 PES 分组包的包头中 stream_id_extension 字段的取值应为 '0x41'；
- 3) 接收端从 PES 分组包中提取主位流，将主位流输入解码器进行解码；

b) 当传输流中包含主位流和知识位流时，传输流的拆解需要解析主位流和知识位流之间的依赖关系。有以下两种方法：

- 1) 接收端解析 PAT 用于定位感兴趣节目的 PMT 所在的传输流分组包，解析包含 library_stream_flag 值为 0 且 library_picture_enable_flag 值为 1 的 PMT，从该 PMT 覆盖的有效范围中的传输流分组包中获取包含主位流 PES 分组包，并解析 library_stream_flag 值为 1 的 PMT，从该 PMT 覆盖的有效范围中的传输流分组包中获取包含知识位流 PES 分组包；
- 2) 接收端解析传输流分组包获得 PES 分组包，如果 PES 分组包的包头中 stream_id_extension 字段的取值为 '0x41'，那么该 PES 分组包中携带的是主位流；如果 PES 分组包的包头中 stream_id_extension 字段的取值为 '0x42'，那么该 PES 分组包中携带的是知识位流。

接收端从知识位流所在的 PES 分组包中提取知识位流，将知识位流输入知识位流解码器进行解码；从主位流所在的 PES 分组包提取主位流，将主位流输入主位流解码器进行解码。

C.4 节目流解封装

AVS3 视频基本流的节目流解封装流程如下：

a) 当节目流中只包含不依赖知识位流的主位流时，节目流的拆解操作如下：

- 1) 接收端在节目流中搜索包含 PSM 的 PES 分组包的起始码，解析包含 `library_stream_flag` 值为 0 且 `library_picture_enable_flag` 值为 0 的 PSM，从该 PSM 覆盖的有效范围中的 PES 分组包中得到主位流；
- 2) 接收端从 PES 分组包中提取主位流，其中这些 PES 分组包的包头中 `stream_id_extension` 字段的取值应为 ‘0x41’，并将主位流输入解码器进行解码。

b) 当节目流中包含主位流和知识位流时，节目流的拆解需要解析主位流和知识位流之间的依赖关系。有以下两种方法：

- 1) 接收端在节目流中搜索包含 PSM 的 PES 分组包的起始码，解析包含 `library_stream_flag` 值为 0 且 `library_picture_enable_flag` 值为 1 的 PSM，从该 PSM 覆盖的有效范围中的 PES 分组包中得到主位流，并解析 `library_stream_flag` 值为 1 的 PSM，从该 PSM 覆盖的有效范围中的 PES 分组包中得到被依赖的知识位流；
- 2) 接收端解析 PES 分组包的包头，如果 PES 分组包的包头中 `stream_id_extension` 字段的取值为 ‘0x41’，那么该 PES 分组包中携带的是主位流；如果 PES 分组包的包头中 `stream_id_extension` 字段的取值为 ‘0x42’，那么该 PES 分组包中携带的是知识位流。

接收端从知识位流所在的 PES 分组包中提取知识位流，将知识位流输入知识位流解码器进行解码；从主位流所在的 PES 分组包中提取主位流，将主位流输入主位流解码器进行解码。