ICS 35.040

CCS L 71

团体标准

T/AI 109.6—XXXX

|  |
| --- |
|  |

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

（征求意见稿）

（在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上）

XXXX - XX - XX实施

中关村视听产业技术创新联盟 发布

XXXX - XX - XX发布

目  次

[目次 I](#_Toc25314)

[前言 II](#_Toc19945)

[引言 III](#_Toc23778)

[信息技术 智能媒体编码 第6部分：智能媒体格式 1](#_Toc31997)

[1 范围 1](#_Toc15878)

[2 规范性引用文件 1](#_Toc15756)

[3 术语和定义 2](#_Toc17156)

[4 缩略语 5](#_Toc690)

[5 AVS3视频文件格式 7](#_Toc32475)

[6 AVS3视频CMAF轨道和媒体配置 15](#_Toc5747)

[7 DASH传输信令 18](#_Toc12892)

[8 SMT传输信令 21](#_Toc25022)

[9 传输流和节目流技术要求 27](#_Toc2226)

[10 RTP传输技术要求 34](#_Toc31034)

[附　录　A （规范性） MIME类型的‘codecs’参数 45](#_Toc10601)

[A.1 基本要求 45](#_Toc15805)

[A.2 AVS3视频编码参数 45](#_Toc21388)

[附　录　B （资料性） 知识位流编码媒体数据传输 46](#_Toc5485)

[B.1 SMT智能媒体传输 46](#_Toc4814)

[B.2 MFU传输 48](#_Toc10956)

[附　录　C （资料性） AVS3视频位流的传输流和节目流封装和解封装 50](#_Toc26332)

[C.1 传输流生成 50](#_Toc18754)

[C.2 节目流生成 51](#_Toc13889)

[C.3 传输流解封装 52](#_Toc27363)

[C.4 节目流解封装 53](#_Toc29070)

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、咪咕文化科技有限公司。

本文件主要起草人：黄成、胡颖、虞露、于化龙、郑建铧、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、候礼志、殷骄阳、黄铁军、高文、张伟民、赵海英、龙仕强、赵璐、李琳、郭佩佩、柳建龙、冯亚楠、聂国梁、王琦。

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。

1. 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。
2. 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。
3. 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。
4. 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。
5. 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。
6. 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、RTP传输技术要求。

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12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 | 相关章、条 |
| 1 | CN202111309701.X | 视频位流封装、解码、访问方法及装置 | 第5章-第9章 |
| 2 | CN2018114887790 | 处理和传输媒体数据的方法和装置 | 第5章-第9章 |
| 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 | 第5章-第9章 |
| 4 | CN202111125693.3 | 传输流和节目流的生成和处理方法及其装置 | 第5章-第9章 |
| 5 | CN2019100028071 | 视频编解码方法与装置 | 第5章-第9章 |
| 6 | CN201910888383.3 | 视频解码方法、视频编码方法、装置、设备及存储介质 | 第5章-第9章 |
| 7 | CN201910503066.5 | 视频解码方法、视频编码方法、装置、设备及存储介质 | 第5章-第9章 |
| 8 | CN201810974952.1 | 支持大跨度相关性信息编码的标识方法及系统 | 第5章-第9章 |
| 9 | CN202311071302.3 | 一种视频码流处理方法、装置、设备及可读存储介质 | 10.1.1、10.1.2、10.1.3、10.1.4 |
| 10 | CN202411155517.8 | 视频位流封装方法、装置、电子设备及程序产品 | 10.1.1、10.1.4.2、10.2 |
| 11 | CN202111309701.X | 视频比特流封装、解码、访问方法及装置 | 第5章-第9章 |
| 12 | CN202310911325.4 | AVS3视频编码码流的传输方法及装置 | 第10章 |

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、RTP传输技术要求。

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。

1. 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 114-2024 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)

IETF RFC 3550 实时应用程序传输协议（A Transport Protocol for Real-Time Applications）

IETF RFC 3551 具有最小控制的音频和视频会议的RTP配置文件（RTP Profile for Audio and Video Conferences with Minimal Control）

IETF RFC 3711 实时传输协议（The Secure Real-time Transport Protocol (SRTP)）

IETF RFC 4566 会话描述协议（SDP: Session Description Protocol）

IETF RFC 4585 基于实时传输控制协议的反馈的扩展RTP配置文件（Extended RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/AVPF)）

IETF RFC 4648 Base16、Base32和Base64数据编码（The Base16, Base32, and Base64 Data Encodings）

IETF RFC 5124 基于实时传输控制协议反馈的扩展安全RTP配置文件（Extended Secure RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/SAVPF)） IETF RFC 7201保护RTP会话的选项（Options for Securing RTP Sessions）IETF RFC 7202保护RTP框架：为什么RTP不要求单一媒体安全解决方案 （Securing the RTP Framework: Why RTP Does Not Mandate a Single Media Security Solution）

IETF RFC 5104 RTP视听配置文件中带反馈的编解码器控制消息 （Codec Control Messages in the RTP Audio-Visual Profile with Feedback (AVPF)）

IETF RFC 5583 会话描述协议（SDP）中的信令媒体解码依赖性 （Signaling Media Decoding Dependency in the Session Description Protocol (SDP) ）

IETF RFC 7202 保护RTP框架：为什么RTP不要求单一媒体安全解决方案 （Securing the RTP Framework: Why RTP Does Not Mandate a Single Media Security Solution）

1. 术语和定义

下列术语和定义适用于本文件。

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。

AVS3视频存取单元 AVS3 video access unit

AVS3视频呈现单元的编码表达。

编码图像 coded picture

T/AI 109.2-2021中规定的一幅图像的编码表示。

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。

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。

初始化片段 initialization segment

包含有媒体流解码所必需元数据的片段。

档次 profile

T/AI 109.2-2021中规定的语法、语义及算法的子集。

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。

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。

轨道 track

ISO/IEC 14496-12:2022标准中一系列相关样本的集合。

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。

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。

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。



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。

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。

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。



RL图像 Reference Library 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。

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。

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。

位流 bitstream

编码图像所形成的二进制数据流。

显示顺序 display order

显示解码图像的顺序。

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。

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。

知识位流 library stream

T/AI 109.2-2021中规定的包含知识图像的位流。

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。

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）。

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。

媒体感知网元 Media-Aware Network Element

RTP流传输经过的网络实体，能够解析RTP负载头或者解析负载内容，并根据获得信息执行特定操作

1. 缩略语

下列缩略语适用于本文件。

|  |  |  |
| --- | --- | --- |
| AVS  AVS3  CMAF  CPB  DASH  DON  DPB  DTS  FCI  FIFO  FIR  HLG  HTTP  ISO BMFF  IETF  IP  LBM  LCU  LFU  LRU  LRAP  MANE  MIME  MFU  MP  MPD  MTU  PA  PAT  PLI  PQ  PES  PMT  PS  PSI  PSM  PTS  PID  RAP  RFC  RPSI  RTP  SAP  SDP  SDR  SMT  SMTP  SLI  STD  TCP  TS  T-STD  UDP  URI  URL  URN  UTC  UUID  XML | 音视频编解码标准  第三代音视频编解码标准  通用媒体格式  编码图像缓冲区  基于HTTP的动态自适应流媒体  解码顺序编号  解码图像缓冲区  解码时间戳  反馈控制指示  先进先出  全帧内请求  混合对数型伽玛  超文本传输协议  ISO基本媒体文件格式  互联网工程任务组  网际协议  知识层数据缓存模型  最大编码单元  最不常使用  最近最少使用  知识随机访问点  媒体感知网元  多用途互联网邮件扩展类型  媒体最小分割单元  媒体呈现  媒体呈现描述  最大传输单元  包访问  节目关联表  图像丢失指示  感知量化  分组化基本流  节目映射表  节目流  节目特定信息  节目流映射  显示时间戳  包标识符  随机访问点  意见征求备忘录  参考图像选择指示  实时传输协议  流访问点  会话描述协议  标准动态范围  智能媒体传输  智能媒体传输协议  片丢失指示  系统目标解码器  传输控制协议  传输流  传输系统目标解码器  用户数据报协议  统一资源标识符  统一资源定位器  统一资源名称  协调世界时  通用唯一标识符  可扩展置标语言 | (audio and video coding standard)  (the third generation AVS standard)  (common media application format)  (coded picture buffer)  (dynamic adaptive streaming over HTTP)  (decoding order number)  (decoded picture buffer)  (decoding time-stamp)  (feedback control indication)  (first input fist output)  (full intra request)  (hybrid log gamma)  (hypertext transfer protocol)  (ISO base media file format)  (internet engineering task force)  (internet protocol)  (library buffer model)  (largest coding unit)  (least frequently used)  (least recently used)  (library random access point)  (media aware network element)  (multipurpose internet mail extensions)  (media fragment unit)  (media presentation)  (media presentation description)  (maximum transmission unit)  (package access)  (program association table)  (picture loss indication)  (perceptual quantisation)  (packetized elementary stream)  (program map table)  (program stream)  (program-specific information)  (program stream map)  (presentation time-stamp)  (packet identifier)  (random access point)  (request for comments)  (reference picture selection indication)  (real-time transport protocol)  (stream access point)  (session description protocol)  (standard dynamic range)  (smart media transport)  (SMT protocol)  (slice loss indication)  (system target decoder)  (transmission control protocol)  (transport kstream)  (transport system target decoder)  (user datagram protocol)  (uniform resource identifier)  (uniform resource locator)  (uniform resource name)  (coordinated universal time)  (universally unique identifier)  (extensible mark-up language) |

1. AVS3视频文件格式
   1. AVS3视频位流和轨道定义
      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。

* + 1. 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，并支持下列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：

1. 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；
2. 知识图像样本群组：用于指示AVS3视频位流中知识图像；
3. 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。
   * 1. 位流结构

AVS3视频编码位流包括两种位流的类型：主位流与知识位流。主位流为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；知识位流为只包含知识图像的位流。其中，知识图像可以被其他位流中的图像参考，而不用于显示输出。

使用上述位流类型，存储AVS3视频内容的位流中应包含序列头，并符合如下约束：

1. 序列头应是位流中包含的AVS3样本（同步样本或者知识随机访问点样本）的一部分；
2. 序列头也可以存储在AVS3视频轨道的样本入口中。
   * 1. AVS3视频轨道类型

用于封装AVS3视频位流的AVS3视频轨道类型包括：

1. AVS3轨道：
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视频编码图像来表示AVS3视频编码位流中的主位流以及可能与其关联的知识位流；
3. 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位流。此时，AVS3轨道可以引用AVS3知识位流轨道；
4. 当AVS3轨道通过'a3lr'轨道引用引用AVS3知识位流轨道时，也称为AVS3主位流轨道；
5. 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位流。AVS3知识位流轨道由AVS3主位流轨道通过'a3lr'轨道引用来引用。
   * 1. 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时域层级的编码图像。AVS3视频主位流中不同时域层级的编码图像可通过以下方式存储：

1. 所有时域层级的编码图像存储在同一个AVS3轨道：
2. AVS3轨道应使用5.3.1中定义的样本入口类型'avs3'；
3. AVS3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；
4. 所有时域层级的编码图像存储在多个不同的AVS3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：
5. 包含最低时域层级的AVS3轨道应使用5.3.1中定义的样本入口类型'avs3'；
6. 不包含最低时域层级的AVS3轨道应使用5.3.2中定义的样本入口类型'lav3'；
7. 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'avs3'的AVS3轨道，并且可选地引用一个或多个样本入口类型为'lav3'的包含较低时域层级的AVS3轨道。通过解析'ltrf'类型的轨道引用，包含不同时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时域层级中的最高时域层级对应的主位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时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时域层级的所有AVS3轨道，重建AVS3视频编码位流中对应时域层级的子位流。

客户端首先确定AVS3视频编码位流中所要解码的最高时域层级，并获取包含该时域层级的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时域层级所需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。

* 1. 样本与解码器配置定义
     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。

* + 1. 解码器配置信息
       1. AVS3解码器配置记录
          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。

* + - * 1. 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;

}

* + - * 1. 语义

sequence\_header\_length：以字节为单位指示配置记录所适用的位流中序列头的长度。

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

1. 如果配置记录适用于主位流，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位流，则序列头中的library\_stream\_flag字段取值‘1’；
2. 如果配置记录适用于主位流，并且该主位流不参考知识位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位流，并且该主位流参考知识位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位流的类型。如果配置记录适用于主位流，library\_dependency\_idc还指示主位流是否参考知识位流中的知识图像进行解码。library\_dependency\_idc的值与描述见表1。

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位流，且该主位流不参考知识位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位流，且该主位流参考知识位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位流 |
| 11 | 当前解码器配置记录所对应的轨道中同时封装主位流和知识位流，且主位流参考知识位流中的知识图像进行解码 |

* + - 1. LAVS3解码器配置记录
         1. 定义

LAVS3解码器配置记录适用于支持时域分层的AVS3视频位流。

* + - * 1. 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(1) reserved = '1'b;  
 unsigned int(18) temporal\_bit\_rate\_lower[i];   
 unsigned int(12) temporal\_bit\_rate\_upper[i];  
 bit(2) reserved = '11'b;

}

}

* + - * 1. 语义

num\_temporal\_layers：指示解码器配置记录所适用的位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位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。

* + - 1. AVS3解码器配置数据盒
         1. 定义

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

* + - * 1. 语法

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

* + - * 1. 语义

Avs3DecoderConfigurationRecord在5.2.2.1.1中定义。

* + - 1. LAVS3解码器配置数据盒
         1. 定义

LAvs3ConfigurationBox包含5.2.2.2.1规定的LAVS3解码器配置记录。

* + - * 1. 语法

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

* + - * 1. 语义

LAvs3DecoderConfigurationRecord在5.2.2.2.1中定义。

* 1. ISO基本媒体文件格式扩展
     1. AVS3样本入口
        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轨道应包含符合T/AI 109.2-2021规定的AVS3视频位流，并符合如下约束：

1. 不包含时域层级的AVS3轨道的每个样本入口应包含一个Avs3ConfigurationBox；
2. 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；
3. 包含时域层级且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。
   * + 1. 语法

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

* + - 1. 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）。

* + 1. 时域分层样本入口
       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

包含时域层级且不包含最低时域层级的AVS3轨道，应使用样本入口类型等于'lav3'的LAvs3SampleEntry。

* + - 1. 语法

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

* + 1. 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'：

1. 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'；
2. SchemeTypeBox存在于RestrictedSchemeInfoBox中，并且scheme\_type设置为'av3l'。
   * 1. 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：

1. 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；
2. 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。
   * 1. 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）。

* + 1. 片段索引数据盒
       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：

1. 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；
2. 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。
   * + 1. 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;  
 }  
}

* + - 1. 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。

* 1. 样本群组定义
     1. LRAP样本群组
        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：

1. LRAP样本只能以知识图像样本作为参考图像样本；
2. 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。

* + - 1. 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;  
 }

}

* + - 1. 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。

* + 1. 知识图像样本群组
       1. 定义

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

* + - 1. 语法

class LibrarySampleGroupEntry extends VisualSampleGroupEntry ('a3lg') {

}

* + 1. 时域层级样本群组
       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。

* + - 1. 语法

class TemporalLayerEntry extend VisualSampleGroupEntry ('telg') {  
 unsigned int(8) temporal\_layer\_id;  
}

* + - 1. 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。

1. AVS3视频CMAF轨道和媒体配置
   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'。

* 1. CMAF轨道约束
     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。

* + 1. 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：

1. 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；
2. 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；
3. 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'）。
   * 1. **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：

1. 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；
2. 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；
3. 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：
4. 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；
5. 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；
6. 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；
7. full\_range\_flag携带T/AI 109.2-2021中定义的样值范围sample\_range字段值。
   * 1. 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。

* + 1. 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：

1. 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：
2. library\_stream\_flag字段的值应设置为‘0’；
3. library\_picture\_enable\_flag字段的值应设置为‘0’；
4. 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：
5. library\_stream\_flag 字段的值应设置为‘0’；
6. 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。

* 1. CMAF切换集约束
     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：

1. 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；
2. 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；
3. 单一初始化CMAF切换集应符合6.3.2中定义的附加约束。
   * 1. 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：

1. 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；
2. 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：
3. 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；
4. 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；
5. 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；
6. 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；
7. 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。
   1. AVS3视频位流约束
      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视频位流，应遵守T/AI 109.2-2021中定义的通用配置和级别语义。

* + 1. 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：

1. 以下字段应具有如下预定值：
2. progressive\_sequence字段的值应等于‘1’；
3. field\_coded\_sequence字段的值应等于‘0’；
4. 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：
5. profile\_id；
6. level\_id。
   1. 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。

1. DASH传输信令
   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。

* 1. DASH片段格式
     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'。

* + 1. 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。

* + 1. 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：

1. 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：
2. library\_stream\_flag=0；
3. library\_picture\_enable\_flag=0；
4. 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：
5. library\_stream\_flag=0；
6. 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。

* + 1. 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：

1. 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；
2. 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。
   1. 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。

* 1. DASH MPD描述子
     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。

* + 1. 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。

* + 1. 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位流类型以及知识位流依赖信息，其中位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。

1. 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>

* + 1. 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。

1. 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：

1. 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：
2. @schemeIdUri=urn:avs:avs3:p6:2022:ColourPrimaries且@value='1'；
3. @schemeIdUri=urn:avs:avs3:p6:2022:MatrixCoefficients且@value='1'；
4. @schemeIdUri=urn:avs:avs3:p6:2022:TransferCharacteristics且@value='1'或'6'；
5. 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：
6. @schemeIdUri=urn:avs:avs3:p6:2022:ColourPrimaries且@value='9'；
7. @schemeIdUri=urn:avs:avs3:p6:2022:MatrixCoefficients且@value='8'或'9'；
8. @schemeIdUri=urn:avs:avs3:p6:2022:TransferCharacteristics且@value='14'；
9. 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：
10. @schemeIdUri=urn:avs:avs3:p6:2022:ColourPrimaries且@value='9'；
11. @schemeIdUri=urn:avs:avs3:p6:2022:MatrixCoefficients且@value='8'或'9'；
12. @schemeIdUri=urn:avs:avs3:p6:2022:TransferCharacteristics且@value='12'。
13. SMT传输信令
    1. 概述

AVS3视频编码位流支持通过T/AI 114-2024 SMT协议进行传输，传输信令基于T/AI 114-2024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。

* 1. 缓存内容更新信令
     1. LBM消息
        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。

* + - 1. 语法

LBM消息的语法见表4。

1. 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** |

* + - 1. 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。

* + 1. LBM反馈消息
       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。

* + - 1. 语法

LBM反馈消息的语法见表5。

1. 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** |

* + - 1. 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 114-2024。

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。

* 1. 包含非对齐MFU依赖关系的样本格式
     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。

* + 1. 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) reserved = 0;

if (reference\_Sample\_flag&&is\_firstMFUinSample\_flag) {

unsigned int(32) depended\_Sample\_id;

}

}

* + 1. 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。

* 1. Asset关系信息描述符
     1. 定义

SMT中定义了Asset关系信息描述符，用于指示同一个SMT Package中Asset的关联关系。使用library\_flag来描述当前Asset与非对齐时间段的知识位流Asset的依赖关系。

* + 1. 语法

Asset关系信息描述符的语法见表6。

1. Asset 关系信息描述符语法

|  |  |  |  |
| --- | --- | --- | --- |
| 语法 | 值 | 位数 | 类型 |
| Asset\_relationship\_information\_descriptor() {  **descriptor\_tag**  **descriptor\_length**  **library\_flag**  **reserved**  **combine\_quality\_flag**  **dependency\_flag**  **composition\_flag**  **equivalence\_flag**  **similarity\_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**  } | ‘11’  N1  N2  N3 | 16  16  1  2  1  1  1  1  1  8  8  8  8  8 | **uimsbf**  **uimsbf**  **blsbf**  **blsbf**  **blsbf**  **blsbf**  **blsbf**  **blsbf**  **blsbf**  **uimsbf**  **uimsbf**  **uimsbf**  **uimsbf**  **uimsbf** |

表6（续）

|  |  |  |  |
| --- | --- | --- | --- |
| 语法 | 值 | 位数 | 类型 |
| }  if(similarity\_flag)  {  **similarity\_selection\_level**  **num\_similarities**  for(i = 0; i <N4; i++) {  asset\_id()  **similarity\_selection\_level**  }  }  if(combine\_quality\_flag) {  **combine\_quality\_ranking**  **num\_combine\_assets**  for(i=0; i<N5;i++){  asset\_id()  }  }  if(library\_flag)  {  **num\_libraries**  for(i = 0; i <N5; i++) {  asset\_id()  }  }  } | N4  N5  N6 | 8  8  8  8  8  8 | **uimsbf**  **uimsbf**  **uimsbf**  **uimsbf**  **uimsbf**  **uimsbf** |

* + 1. 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。

combine\_quality\_flag：该字段为1位。取值为1时表示具备组合关系的多个Asset整体具备一个联合质量等级；取值为0时表示具备组合关系的多个Asset不存在联合质量等级。

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。值为‘0’表示不需要添加，值为‘1’表示需要添加。

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被缺省呈现。当缺省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被缺省呈现。当缺省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\_combine\_assets：该字段为8位。指示与此描述符所描述的Asset有联合质量等级关系的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关系信息描述符中的asset\_id：

1. 当Asset关系信息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；
2. 当Asset关系信息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；
3. 当Asset关系信息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；
4. 当Asset关系信息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；
5. 当Asset关系信息描述符用于指示联合质量关系时，asset\_id字段指示与此描述符所描述的Asset有联合质量关系关系的Asset的标识符；
6. 当Asset关系信息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 114-2024。

* 1. 彩色信息描述符
     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。

* + 1. 语法

AVS3彩色信息描述符的语法见表7。

1. 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** |

* + 1. 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 114-2024。

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。

1. 传输流和节目流技术要求
   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视频位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：

1. AVS3视频流应是ISO/IEC 13818-1:2022中节目的一个节目元素，位流的stream\_type字段值应等于‘0xD4’；
2. 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位流中；
3. 在stream\_type等于‘0xD4’且具有ISO/IEC 13818-1:2022中层次描述符的位流中，层次描述符中的hierarchy\_type字段值应等于‘3’（时域可扩展性）；
4. 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位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：

1. 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；
2. 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。
   1. PES分组
      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：

1. 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；
2. 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。

* + 1. 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：

1. 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；
2. 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；
3. 当data\_alignment\_indicator值为‘0’时，表示没有定义是否有任何对齐方式。
   * 1. 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。

* 1. 节目和节目元素描述符
     1. 节目和节目元素描述子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。

1. 节目和节目元素描述符

|  |  |  |  |
| --- | --- | --- | --- |
| 描述符标签值 | TS | PS | 标识 |
| 6 | X | X | data\_stream\_alignment\_descriptor |
| 209 | X | X | AVS3\_video\_descriptor |

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。

* + 1. 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位流中的AVS3视频最高时域层级的信息。

1. 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** |
| } |  |  |

* + 1. 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。

1. 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  100.0  119.88  120.0 | 23.976 24.0 29.97 30.0 59.94  50.0  59.94  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位流是否为知识位流。值为‘1’表示该位流是知识位流；值为‘0’表示该位流是主位流。

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。

* + 1. 数据流对齐描述符

ISO/IEC 13818-1:2022中规定的数据流对齐描述符描述了AVS3视频位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。

1. 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 |

* + 1. 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视频位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。

* 1. 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：

1. 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；
2. 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。

* 1. 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。

t(i)

传输流解复用

Demux

其他位流

TBn

…

MBn

EBn

D

DPB

Rxn

Rbxn

An(j), tdn (j)

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。

TBn——位流n的传输缓存。

TBS——传输缓存TBn的大小，以字节为单位。

MBn——位流n的复用缓存。

MBSn——复用缓存MBn的大小，以字节为单位。

EBn——AVS3视频编码位流的位流缓存。

j——AVS3视频编码位流中AVS3视频存取单元的索引。

An(j)——AVS3视频编码位流的第j个存取单元。

tdn(j)——An(j)在系统目标解码器中的解码时间，以秒为单位。

Rxn——从传输缓存TBn到复用缓存MBn的传输速率。

Rbxn——从复用缓存MBn到位流缓存EBn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。

* + 1. 缓存管理

缓存管理符合下列约束：

1. 仅有一个传输缓存TBn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；
2. 仅有一个复用缓存MBn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MBSn约束见公式（1）：

MBSn = BSmux + BSoh + MaxCPB[level\_id] – BBS ………………（1）

式中：

MBSn——复用缓存，单位为位；

BSoh——包头缓存，单位为位，定义见公式（2）：

BSoh = (1/750) × max{MaxBR[level\_id], 2 000 000} ………………（2）

式中：

BSmux——额外复用缓存，单位为位，定义见公式（3）：

BSmux = 0.004 × max{MaxBR[level\_id], 2 000 000 } ………………（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。

1. 仅有一个位流缓存EBn用于接收分层描述子对应的位流集合中的所有位流，该位流缓存的总大小EBSn定义见公式（4）：

EBSn= BBS ………………（4）

式中：

EBSn——位流缓存的总大小，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。

1. 从TBn到MBn的传输应按照如下方式执行：

当TBn中没有数据时，Rxn设为0；否则，计算方式见公式（5）：

Rxn = BitRate × 400 ………………（5）

式中：

Rxn——从传输缓存TBn到复用缓存MBn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。

1. 从MBn到EBn的传输应按照公式（6）执行：

Rbxn = MaxBR[level\_id] ………………（6）

式中：

Rbxn——从复用缓存MBn到位流缓存EBn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MBn中存在PES分组包负载数据且EBn没有充满，PES分组包负载以速率Rbxn从MBn被传输到EBn。如果EBn已经被充满，MBn中的数据不应被移除。当一个字节的数据从MBn被传输到EBn时，MBn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MBn中不存在PES分组包负载数据时，没有数据被从MBn中移除。所有进入MBn的数据都应从其中离开。所有离开MBn的PES分组包负载数据都要立即进入EBn。

* + 1. STD延时

STD延时是数据通过STD中的TBn、MBn和EBn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An(j)中的所有字节i，tdn(j)–t(i)≤10（s）。

* + 1. 缓存管理条件

传输流的构建应符合如下的缓存管理条件：

1. 每个TBn不应上溢，且应在每一秒中至少有一次处于被清空的状态；
2. 每个MBn、EBn和DPB不应上溢；
3. EBn不应下溢。当AVS3视频存取单元An(j)中一个或多个字节在解码时间tdn(j)不存在于EBn中时，EBn对于An(j)发生下溢。
4. RTP传输技术要求
   1. AVS3 视频的RTP 负载格式
      1. 概述

AVS3视频的RTP负载以AVS3视频元码流形式进行封装。一个AVS3视频的RTP负载中可以包含一个或者多个AVS3视频元码流，或一个AVS3视频元码流分片。AVS3视频元码流为符合T/AI 109.2-2021的视频编码位流片段，由位流中相邻两个起始码或相邻若干个起始码之间的编码数据组成（包含第一个起始码），根据起始码取值识别AVS3视频元码流数据类型，如序列头、用户数据、扩展数据、帧内/帧间预测图像。其中，帧内/帧间预测图像类型的AVS3视频元码流，由帧内/帧间预测图像起始码（包含）开始到下一个帧内/帧间预测图像起始码、序列起始码、序列结束码、或视频编辑码之前的连续字节组成。AVS3视频元码流分片为将AVS3视频元码流以字节为最小单位进行分割形成的连续片段。

AVS3元码流的起始码是一组特定的位串。在符合T/AI 109.2-2021的位流中，除起始码外的任何情况下都不应出现这些位串。起始码由起始码前缀和起始码值构成。起始码前缀是位串‘0000 0000 0000 0000 0000 0001’；起始码值是一个8位整数，用来表示负载数据类型。所有的起始码都应字节对齐。

本章节定义的所有整型字段均遵循RTP定义的网络字节序，即高位字节优先。

* + 1. RTP头的使用

一个AVS3视频的RTP数据包包含RTP头和AVS3视频负载两部分，如图2所示。

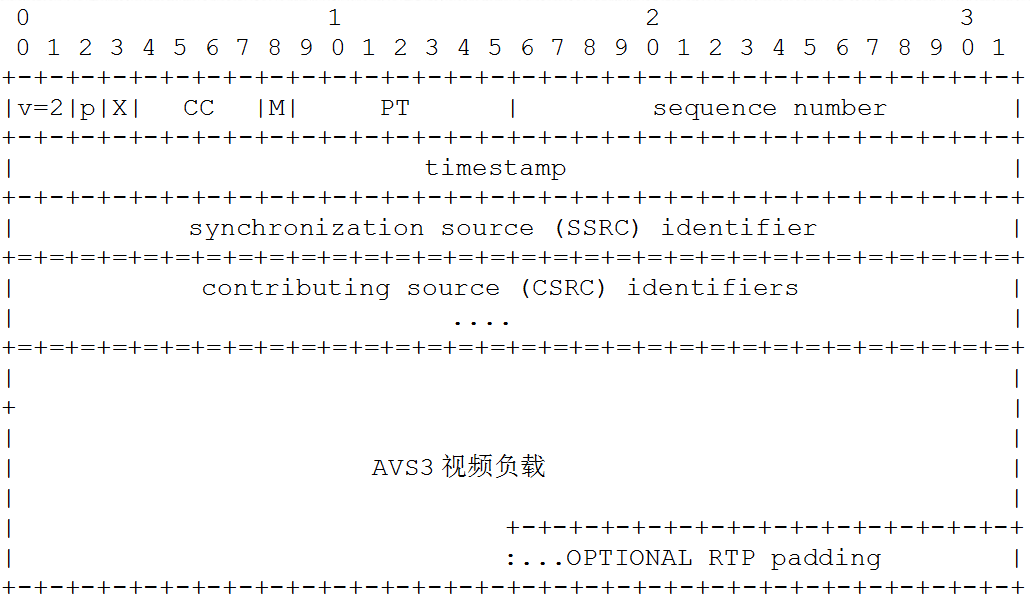


图2 AVS3视频的RTP数据包

IETF RFC 3550中定义了RTP头的结构，根据AVS3视频的RTP数据包格式，RTP头中的字段应设置如下：

Marker bit （M）：该字段为1位，指定视频帧的边界，视频帧的最后一个RTP数据包的标记位设为1，其他RTP数据包标记位均设为0。

payload type（PT）：该字段为7位，根据IETF RFC 3551的定义，上层应用在每个RTP会话中（如SDP中）应动态分配一个PT值并维护PT值与负载格式的映射关系。PT值的取值范围应遵循IETF RFC 3551规定（例如96-127）。

Timestamp：该字段为32位，RTP时间戳设为编码内容的采样时间戳，所用的时钟为90000Hz，即RTP时间戳的时间单位为1/90000s。时间戳的初始值应该是随机的，同一帧图像的不同RTP数据包的时间戳应该一致。对于序列头、用户扩展数据、图像头不具备时间属性的负载，其时间戳应与紧随其后的图像数据的时间戳一致。

Sequence number：该字段为16位，表示RTP数据包的传输序列号，每发送一个RTP数据包，序列号加1，接收方可以用它来检测数据包丢失和恢复数据包序列。序列号的初始值应该是随机的。

RTP头中其它字段（V、P、X、CC等）应遵循IETF RFC 3550的规定。

* + 1. AVS3视频的RTP负载头

##### 10.1.3.1定义

AVS3视频的RTP数据包中的AVS3视频负载结构由AVS3视频的负载头和AVS3视频位流数据组成。其中，AVS3视频负载头包含一个AVS3视频通用负载头、一个AVS3视频扩展负载头（可选）和一个或多个负载结构对应的负载头。针对负载结构类型，负载结构对应的负载头可以为单一负载头、分片负载头、聚合负载头。

##### 10.1.3.2语法

AVS3视频负载的第一个字节为AVS3视频通用负载头，该字节为固定的，表示负载的AVS3视频位流数据的基本信息，语法结构见图3。

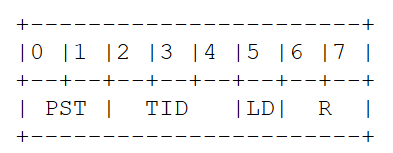


图3 AVS3视频通用负载头

##### 10.1.3.3语义

PST：表示payload structure type，该字段为2位，指示RTP负载类型，取值0表示单一负载结构（见10.1.4.2）；取值1表示分片负载结构(见10.1.4.3)；取值2表示聚合负载结构(见10.1.4.4)；取值3保留。

TID：表示temporal id，该字段为3位，无符号整数，指示负载所属的时间层标识，取值范围是0~7，其取值应与T/AI 109.2-2021中的temporal\_id的值相同。

LD：表示library dependency，该字段为1位，无符号整数，指示负载所属的位流类型，包括主位流和知识位流。取值为0时指示当前负载为AVS3视频主位流的元码流数据；取值为1时指示当前负载为AVS3视频知识位流的元码流数据。

R：表示保留位，该字段为2位，无符号整数。

* + 1. RTP负载结构

##### 10.1.4.1概述

RTP负载的AVS3视频位流数据可以是单个元码流数据、单个元码流分片数据以及多个元码流数据，分别对应单一负载、分片负载和聚合负载的RTP负载结构类型。对AVS3视频元码流进行RTP封装时，需考虑IP层MTU的限制，选择不同的负载结构，见10.2封包规则。

10.1.4.2单一负载结构

###### 10.1.4.2.1定义

在单一负载结构下RTP负载只包含一个AVS3视频元码流，并且通用负载头中的参数payload structure type取值为0。单一负载中，AVS3视频负载包括AVS3视频通用负载头、单一负载头以及一个AVS3视频元码流。

当SDP参数sprop-max-don-diff取值大于0时，表示RTP数据包的解码顺序与传输顺序不一致，必须在AVS3视频元码流数据前增加2个字节无符号整数参数DON（Decoding Order Number）用于表示负载的解码顺序。DON的起始值为随机数，且是以1为单位递增的连续整数，到达最大值后从0开始重新编号。

单一负载结构的RTP负载数据的结构示例如图4所示。

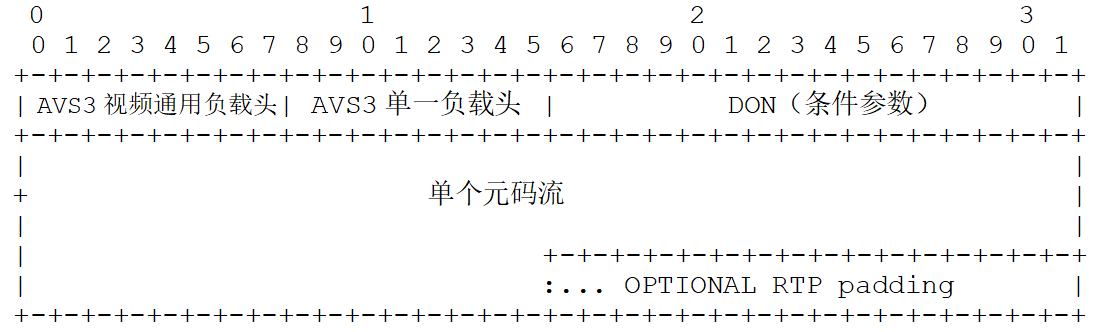


图4单一负载结构下的RTP负载数据的结构示例

###### 10.1.4.2.2语法

AVS3视频的单一负载头长度为1个字节，用于描述单一负载结构下负载的AVS3视频元码流的具体信息，语法结构见图5。

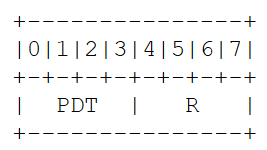


图5 单一负载头

###### 10.1.4.2.3语义

PDT：表示payload data type，该字段为4位，无符号整数，指示负载的数据类型，其取值及类型说明见表12。

表12 AVS3视频负载数据类型定义

|  |  |  |
| --- | --- | --- |
| 取值 | 类型说明 | 数据内容 |
| 0 | 序列头 | 起始码值为B0，元码流包含序列头数据 |
| 1 | 序列头后视频扩展数据 | 起始码值为 B5 ，元码流包含序列头后的视频扩展数据。 |
| 2 | 序列头后用户数据 | 起始码值为 B2，元码流包含序列头后的用户数据。 |
| 3 | I 帧 | 起始码值为 B3，元码流包含I帧的图像头及其图像头后视频扩展数据、用户数据和所有片数据。 |
| 4 | RL帧 | 起始码值为B6，元码流包含RL帧的图像头、图像头后视频扩展数据、图像头后用户数据和所有片数据。 |
| 5 | P帧 | 起始码值为B6，元码流包含P帧的图像头、图像头后视频扩展数据、图像头后用户数据和所有片数据。 |
| 6 | B帧 | 起始码值为B6，元码流包含B帧的图像头、图像头后视频扩展数据、图像头后用户数据和所有片数据。 |
| 7 | 序列结束码 | 起始码值B1，元码流包含序列结束码。 |
| 8 | 视频编辑码 | 起始码值为B7，元码流包含视频编辑码。 |
| 9~15 | 保留，未定义 | |

R：表示保留位，该字段为4位，无符号整数。

10.1.4.3分片负载结构

###### 10.1.4.3.1定义

在分片负载结构下RTP负载只包含一个AVS3视频元码流分片，并且通用负载头中的参数payload structure type取值为1。根据IP层MTU的限制，需将元码流分割成合适大小的AVS3视频元码流分片再进行RTP封装。同一个AVS3视频元码流的元码流分片，按位流顺序分割，并按传输序列号的连续顺序发送，即，同一RTP流中的其他数据包不得在该AVS3视频元码流的第一个分片和最后一个分片之间发送。

在分片负载结构下，AVS3视频负载包括AVS3视频通用负载头、分片负载头以及一个AVS3视频元码流分片。

当sprop-max-don-diff取值大于0且fragment\_start =1时，必须在元码流分片前增加2个字节无符号整数参数DON用于表示负载的解码顺序。DON的起始值为随机数，且是以1为单位递增的连续整数，到达最大值后从0开始重新编号。

分片负载结构的RTP负载数据的结构示例如图6所示。

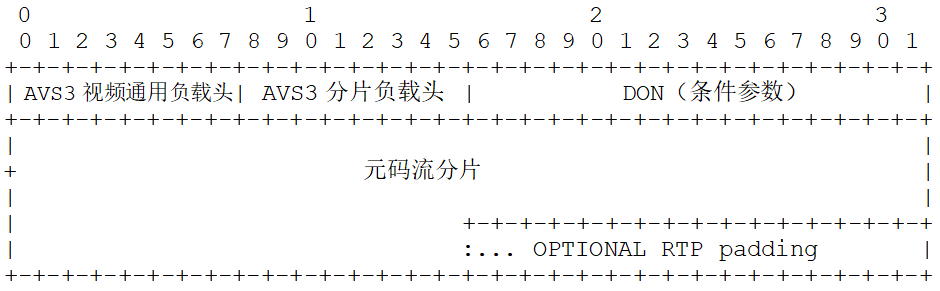


图6 分片负载结构下的RTP负载数据的结构示例

###### 10.1.4.3.2语法

AVS3分片负载头长度为1个字节，用于描述分片负载结构下负载的AVS3视频元码流的具体信息，语法结构见图7。

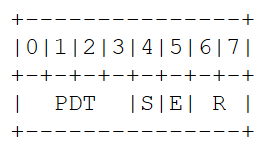


图7 分片负载头

###### 10.1.4.3.3语义

PDT：表示payload data type，该字段为4位，无符号整数，指示分片负载所属的元码流数据类型，其取值及类型说明见表12。

S：表示fragment start，该字段为1位，取值为1时表示负载是元码流的起始切片；取值为0时表示负载不是元码流的起始切片。

E：表示fragment end，该字段为1位，取值为1时表示负载是元码流的结束切片；取值为0时表示负载不是元码流的结束切片。

注：一个RTP数据包中，fragment start和fragment end不能同时为1。

R：表示保留位，该字段为2位，无符号整数。

10.1.4.4聚合负载结构

###### 10.1.4.4.1定义

在聚合负载结构下RTP负载包含多个（至少2个）AVS3视频元码流的数据，并且通用负载头中的参数payload structure type取值为2。聚合负载结构下，具体的负载数据中不可以包含AVS3视频元码流分片。

在聚合负载结构下，AVS3视频负载包括AVS3通用负载头、多个聚合负载头以及对应的多个AVS3视频元码流，其中，每个AVS3视频元码流前应包含一个标识元码流尺寸大小的字段（size），长度为2个字节无符号整数。

当sprop-max-don-diff取值大于0，在每个元码流的size前，AVS3聚合负载头后必须增加2个字节无符号整数参数DON用于表示负载的解码顺序。DON的起始值为随机数，且是以1为单位递增的连续整数，到达最大值后从0开始重新编号。

聚合负载结构的RTP负载数据的结构示例如图8所示。

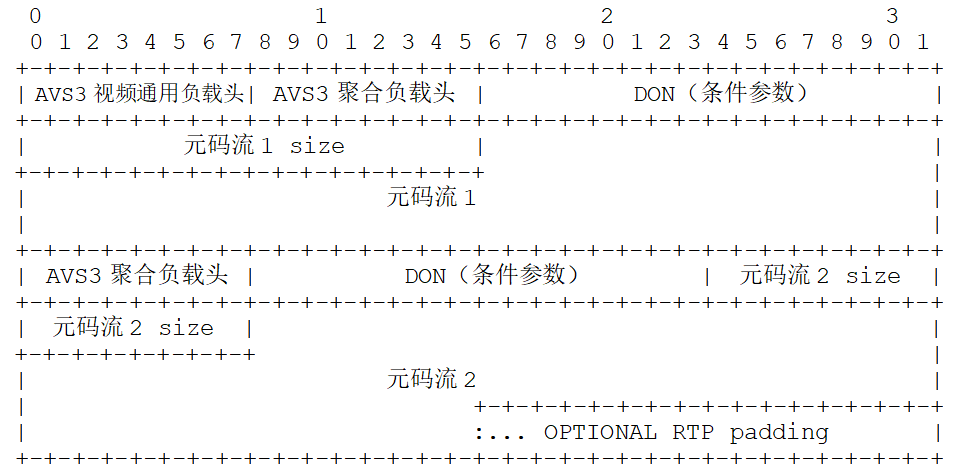


图8 聚合负载结构下的RTP负载数据的结构示例

###### 10.1.4.4.2语法

AVS3聚合负载头长度为1个字节，用于描述聚合负载结构下负载的每个AVS3视频元码流的具体信息，语法结构见图9。

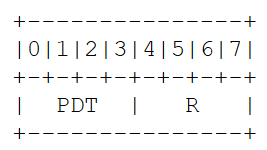


图9 聚合负载头

###### 10.1.4.4.3语义

PDT：表示payload data type，该字段为4位，无符号整数，指示负载所属的数据类型，其取值及类型说明见表12。

R：表示保留位，该字段为4位，无符号整数。

* 1. 封包规则

根据AVS3视频的RTP数据包的负载类型，AVS3视频编码位流数据进行封装时应该满足以下规则：

1. 不同类型的元码流数据不允许聚合在同一个RTP数据包中，以下情况例外：在聚合包尺寸大小满足IP层的MTU大小的限制下，连续的序列头、序列头后视频扩展数据、序列头后用户数据元码流应封装到一个聚合包中。
2. 多帧图像不允许聚合在同一个RTP数据包中。
3. 当一个AVS3视频元码流封装成IP包尺寸超过IP层MTU大小，必须将元码流进行分割，并采用分片负载格式对分割后的元码流分片进行封包传输。
4. 视频序列结束码或者视频编辑码类型的元码流应封装为单包。
5. 聚合包和分片包不可以相互嵌套，即聚合负载中不可以包含元码流分片。
6. 非独立主位流与知识位流的元码流数据不允许聚合在同一个RTP数据包中。
7. RTP 数据包的负载数据不允许嵌套，即RTP负载中不可以包含完整的RTP数据包。
   1. 解包流程

解包过程就是将AVS3视频编码位流从接收的多个RTP数据包中解析出来，并且按照解码顺序传送给解码器。

当AVS3视频编码位流的传输顺序与解码顺序一致时，接收端根据RTP头中的Sequence number顺序进行组帧后，按顺序传递给解码器进行解码。当AVS3视频编码位流的传输顺序与解码顺序不一致，接收端根据RTP头中的Sequence number顺序进行组帧后，并且以DON的顺序传递给解码器进行解码。

* 1. AVS3视频的RTP负载格式参数
     1. 媒体类型注册

媒体名字：video

媒体子类名字：AVS3

必选参数：无。

可选参数：profile-id、level\_id、sprop-sequence-header、bitstream\_info\_idc、sprop-max-don-diff

* + 1. SDP可选参数说明

在SDP声明性的会话描述中，参数 profile-id、level-id用于声明传输视频流的档次、级别。参数sprop-sequence-header用于指示传输位流的特性；当主位流和知识位流通过不同RTP流传输时，bitstream\_info\_idc用于表示发送端的位流类型；参数sprop-max-don-diff用于指示元码流的传输顺序与解码顺序不一致时的处理方式。

**profile-id：**

profile-id的取值范围应与T/AI 109.2-2021 B.2保持一致。当该参数未给定时，则该参数值缺省为0x20，表示基准8位档次（Main 8bit profile）。

**level-id：**

level-id的取值范围应与T/AI 109.2-2021 B.3保持一致。当该参数未给定时，则该参数值缺省为0x20，表示4.0.30级别。

**sprop-sequence-header：**

sprop-sequence-header表示传输的视频流的序列头，用来描述视频流的属性，比如：传输的视频流的宽、高、帧率、码率等。其值为T/AI 109.2-2021 7.1.2.2定义的视频序列头码流的Base64（参见IETF RFC 4648）编码。

**bitstream\_info\_idc：**

bitstream\_info\_idc用来表示当前发送的AVS3视频位流类型。该字段取值含义如下：

* 0：传输视频流为独立主位流；
* 1：传输视频流为非独立主位流；
* 2：传输视频流为知识位流；
* 3：传输视频流为知识位流和非独立主位流；

当该参数未给定时，则该参数值缺省为0。

**sprop-max-don-diff：**

如果元码流的传输顺序与解码顺序相同，则该参数的值必须等于0。否则，该参数将指定任意两个元码流A和元码流B的解码顺序号（例如decode\_order\_index）值之间的最大绝对差值，其中元码流A在解码顺序上位于元码流B之后，而元码流A在传输顺序上位于元码流B之前。

sprop-max-don-diff 的值必须是一个整数，范围为0至32767（含）范围内的整数。当该参数未给定时，则该参数值缺省为0。

* + 1. 媒体类型参数与SDP域的匹配

SDP文档请参考IETF RFC4566。媒体类型video/AVS3 (针对T/AI 109.2-2021 规定的视频)字符串到SDP域的匹配方法为：

* **"m="**行的媒体名称（media name），对应MIME媒体名字为video
* **"a=rtpmap"**行的编码名称（encoding name），对应MIME子类名称为AVS3
* **"a=rtpmap"**的时钟率参数为 90000

可选参数profile-id、level-id、sprop-sequence-header应在"a=fmtp"行，这些参数的表示形式为一系列由半角分号“;”分开的“参数=值”对。

示例：

m=video 49170 RTP/AVPF 98

a=rtpmap:98 AVS3/90000

a=fmtp:98 profile-id=20; level-id=20;

* + 1. 在 SDP 发起（Offer）/应答(Answer)模型中的用法

当使用RTP来传输T/AI 109.2-2021 视频，并且在协商过程中使用SDP的发起（Offer）/应答（Answer）模型时，应遵守以下规则和限制：

1. 用于识别 T/AI 109.2-2021 视频媒体档次级别的profile-id、level-id参数，如果不填，则分别为缺省值0x20，0x20。
2. 用于识别 T/AI 109.2-2021 视频流属性的序列头sprop-sequence-header参数，Offer/Answer可以不填。
   * 1. 在声明性的会话描述中的用法

当使用RTP来传输T/AI 109.2-2021 视频以声明式方式使用SDP时，比如使用实时流协议（RTSP）或会话声明协议（SAP）传输T/AI 109.2-2021 视频，应遵守以下规则和限制：

1. 参数 profile-id、level-id 、sprop-sequence-header用于声明发送端发送的T/AI 109.2-2021 视频流的档次、级别和视频流属性，而不是接收端需要支持码流的档次、级别和视频流属性。
2. SDP的接收端必须支持SDP中所列举的所有参数的相应取值。否则，接收端应拒绝加入回话。会话的创建者应使用接收应用程序所支持的参数及取值。
   * 1. 多流模式下的依赖性信令

当非独立主位流和知识位流通过同一个会话(session)的不同RTP流传输、或者不同会话的不同RTP流传输时，不同RTP流之间在SDP中的依赖关系依据IETF RFC 5583来描述。会话依赖性的顺序是从包含知识位流的 RTP 流到包含非独立主位流的RTP流。

当AVS3视频的不同时域层级视频位流通过同一个会话的不同RTP流传输、或者不同会话的不同RTP流传输，不同RTP流之间在SDP中的依赖关系依据IETF RFC 5583来描述，会话依赖性的顺序是从包含最低时序子层的RTP 流到包含最高时序子层的RTP流。

* 1. 使用消息反馈
     1. 概述

本节定义了图像丢失指示（PLI）、片丢失指示（SLI）、参考图像选择指示（RPSI）和全帧内请求（FIR）反馈信息在传输AVS3视频位流中的使用。PLI、SLI 和 RPSI 信息在 IETF RFC 4585中定义，FIR 信息在IETF RFC 5104中定义。

* + 1. 图像丢失指示 (PLI)

媒体发送方接收到 PLI时，表示丢失了属于一幅或多幅图像的未定义数量的编码视频数据。对于AVS3视频独立主位流或者采用不同RTP流传输的知识位流和非独立主位流，AVS3视频位流发送方接收到PLI后应发送序列头后的第一个I帧和序列头及其它的参数信息（如用户扩展信息）；然而，有时在媒体接收方已知关于比特流结构的信息情况下，例如，SDP会话时携带了T/AI 109.2-2021定义的序列头sprop-sequence-header的数据，并在会话期间保持不变，则媒体发送方接收到PLI时不必重复发送序列头。

当AVS3视频的知识位流和非独立主位流采用同一个RTP流传输时，AVS3视频位流发送方接收到PLI后，建议发送知识位流和非独立主位流的序列头后的第一个I帧和序列头及其它的参数信息。

注：在任何情况下，都必须遵守 IETF RFC 4585的定时和拥塞控制机制。

* + 1. 片丢失指示 (SLI)

SLI可用来向媒体发送方指示在一幅图像LCU的光栅扫描顺序中丢失了多个LCU。在IETF RFC 4585中定义的 SLI 的反馈控制指示 (FCI) 字段中，字段First必须设置为第一个丢失的LCU地址。由于该地址是按图像LCU的光栅扫描顺序排列的，对于片（patch）中连续的编码单元丢失，需要计算LcuIndex,其中LcuIndex = LcuRow \* PictureWidthInLcu + LcuColumn。当片的尺寸小于图像尺寸且存在多行最大编码单元的数据丢失时，需要拆分成多个SLI的信息发送。

IETF RFC 4585中定义的字段Number必须设置为连续丢失的编码宏块的数目，同样按图像LCU的光栅扫描顺序排列。其中， 字段First和字段Number都是按图像LCU的光栅扫描顺序而不是LCU的比特流顺序计算的。

IETF RFC 4585中定义的字段PictureID必须设置为T/AI 109.2-2021中定义的二进制表示 decode\_order\_index 值的 6 个最小有效位，该值表示丢失LCU的图像帧的解码顺序索引值。

注：AVS3视频传输过程中，SLI信息不支持跨序列边界使用。在AVS3中，宏块大小并非固定的16x16亮度样本，而是可变的。因为最大编码单元LCU的大小设置是序列级功能，并且没有跨序列边界的预测，跨序列边界使用切片丢失指示是没有意义的。

* + 1. 参考图像选择指示 (RPSI)

RPSI用于阻止编码误差传播，以提高抗错能力。解码器侧跟踪已解码图像中的错误，并通知编码器侧图像缓冲区中最近的正确解码的特定图像，则解码器侧应使用RPSI反馈信息，要求编码器在编码下一张图像时使用该正确的图像可用性信息，以阻止进一步的时间误差传播。

如果AVS3视频位流包含知识图像，则AVS3编码器可以使用知识图像来实现上述目的。RPSI反馈信息必须仅在解码器侧存在解码错误图像时使用，即RPSI仅用作参考图片选择请求，以便也能在组播中使用。

其中，IETF RFC 4858定义的Native RPSI bit string字段等于T/AI 109.2-2021中定义的图像头中的解码顺序索引（decode\_order\_index） 字段，用于RPSI反馈信息所指示的图片。

* + 1. 全帧内请求(FIR)

FIR消息用于使编码器尽快发送独立的解码刷新点(例如，遵守IETF RFC 5104中设置的与拥塞控制相关的约束)。

媒体发送方收到FIR后，对于AVS3视频独立主位流或者采用不同RTP流传输的知识位流和非独立主位流，必须发送序列头后的第一个I帧和序列头及其它的参数信息（如用户扩展信息）。然而，有时在媒体接收方已知关于比特流结构的信息情况下，例如，SDP会话时携带了序列头sprop-sequence-header的数据，并在会话期间保持不变，则不必重复发送序列头。

当AVS3视频的知识位流和非独立主位流采用同一个RTP流传输时，AVS3视频流发送方接收到FIR后，建议发送知识位流和非独立主位流的序列头后的第一个I帧和序列头及其它的参数信息。

* 1. 安全考虑（信息性）

采用AVS3视频负载格式的RTP报文遵循IETF RFC 3550和IETF RFC 3551、IETF RFC 4585、IETF RFC 3711、IETF RFC 5124中的安全考虑相关定义。但是如同IETF RFC 7202中提到的一样，采用何种安全解决方案来实现报文的机密性、完整性和源真实性等安全目标，并不是RTP负载格式的职责，而是应用开发者应考虑的事情。应用开发者在实现RTP安全时，可参考IETF RFC 7201所述的安全内容。

相比其他通用的RTP负载格式，除用户数据报文以外，AVS3视频负载格式并没有引入其他安全风险。当RTP报文没有采取安全保护措施时，攻击者可通过注入恶意数据或者仿造报文的方式，诱发接收端的安全风险。例如攻击者在用户数据中插入恶意数据，从而导致接收端处理异常，影响视频播放；又如攻击者构造新的报文，大量耗费接收端解码资源，导致接收端无法服务。因此，在不可靠环境下，应用开发者可以采用数据源认证和数据完整性机制保证报文安全，安全实现机制可参见IETF RFC 3711。

对报文进行安全保护可以防止潜在的安全风险，但也会带来一些负面影响。当RTP报文采用加密传输时，媒体感知网元（MANE）将无法执行基于媒体感知的操作。如MANE支持基于媒体感知的转发，对不同类型帧赋予不同转发优先级，确保拥塞情况下高优先级帧优先转发，如果报文采用加密传输，则该机制无法生效。如果要在加密传输情况下允许执行基于媒体感知的操作，MANE需安全可信且参与了安全上下文的创建。

* 1. 拥塞控制（信息性）

采用AVS3视频负载格式的RTP报文遵循标准IETF RFC 3550和IETF RFC 3551、IETF RFC 4585、IETF RFC 3711、IETF RFC 5124的拥塞控制相关定义。

实时编码场景可通过实时调整码率来实现拥塞控制。如编码器可以利用动态调整QP参数等方式满足传输速率要求。

预先编码场景可通过切换码率文件或者通过裁剪帧方式实现拥塞控制。切换码率文件机制在发送端保存多个不同码率文件，发送端可根据拥塞程度选择不同码率文件发送；裁剪帧方式涉及AVS3的时域分层编码技术，发送端删除时域层级（参考负载头中的TID）较高的编码数据，直到发送比特率下降到可接受的范围。

RTP报文传输遵循协商的档次和级别约定，如果拥塞控制机制导致档次和级别变化，发送端和接收端需重新协商。

MANE也能够参与拥塞控制。如网络拥塞时MANE识别出非参考帧主动丢弃减免拥塞；当检测到属于同一帧的包已经丢弃时，MANE丢弃属于同一帧的所有后续包，以减轻网络负载。

当RTP报文采用加密传输时，除非MANE安全可信且参与了安全上下文的创建，否则MANE无法完成基于媒体感知的丢弃操作。

1. （规范性）  
   MIME类型的‘codecs’参数
   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。

* 1. 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。

1. 档次（profile\_id）：‘codecs’参数的第二个子参数，编码为十六进制数；
2. 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。

1. （资料性）  
   知识位流编码媒体数据传输
   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

* 1. 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：

1. 一个主位流媒体资源；
2. 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：

1. 如果被传输过且还可用则不再重复传输；
2. 如果被传输过但不可用则再次传输；
3. 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

1. （资料性）  
   AVS3视频位流的传输流和节目流封装和解封装
   1. 传输流生成

将AVS3视频编码的位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。

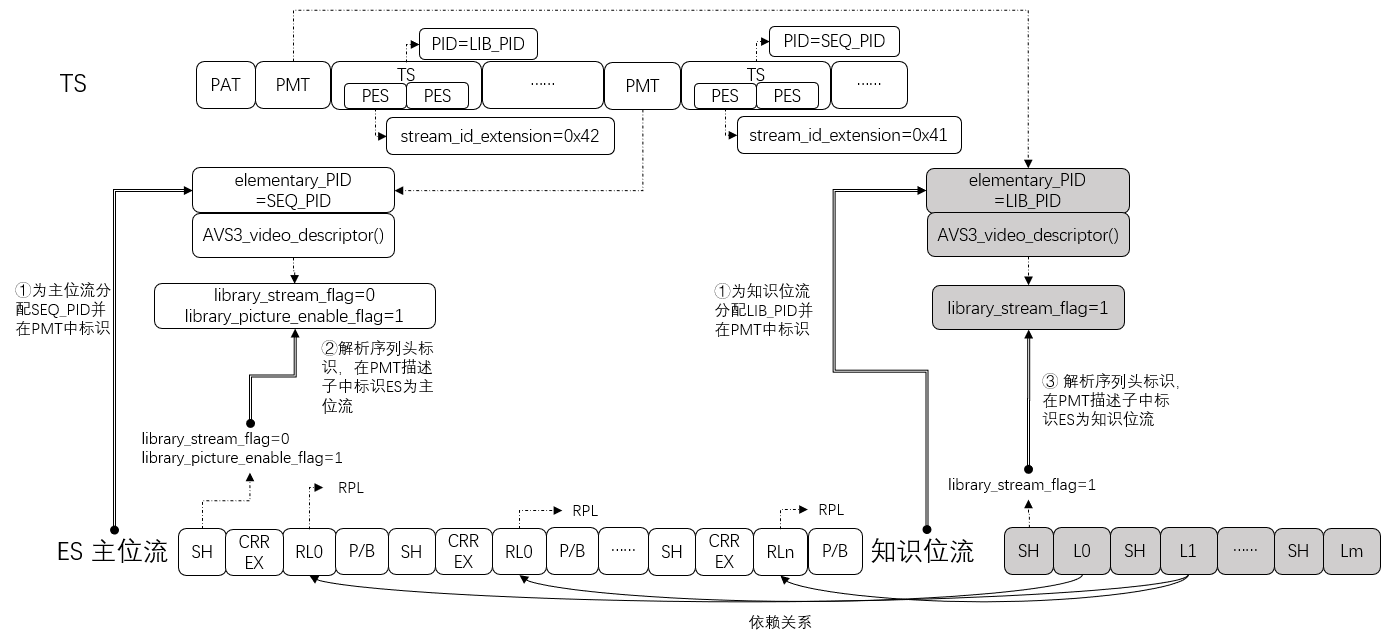
生成传输流应从视频位流中解析语法元素用于为传输流中PSI表中的相关字段设定该字段应取的值。在为视频位流所属的节目建立PMT时，从视频位流的序列头中解析library\_stream\_flag和library\_picture\_enable\_flag的值，方法如下：

1. 当library\_stream\_flag的值为0且library\_picture\_enable\_flag的值为0时，判断该视频位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；
2. 当library\_stream\_flag的值为0且library\_picture\_enable\_flag的值为1时，判断该视频位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；
3. 当library\_stream\_flag的值为1时，判断该视频位流包含知识位流，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视频位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位流，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’。两个位流是一个节目中的节目元素，该节目被分配唯一的program\_number。承载每个位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。两个位流的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。



注：其中虚线从语法元素集指向其包含的语法元素，实线从被依赖的知识图像指向依赖该知识图像的RL图像，双实线表示根据ES的语法元素在TS中的字段赋值的操作。SH表示序列头，CRREX表示cross\_random\_access\_point\_reference\_extension()。

图C.1 传输流封装流程

* 1. 节目流生成

将AVS3视频编码的位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。

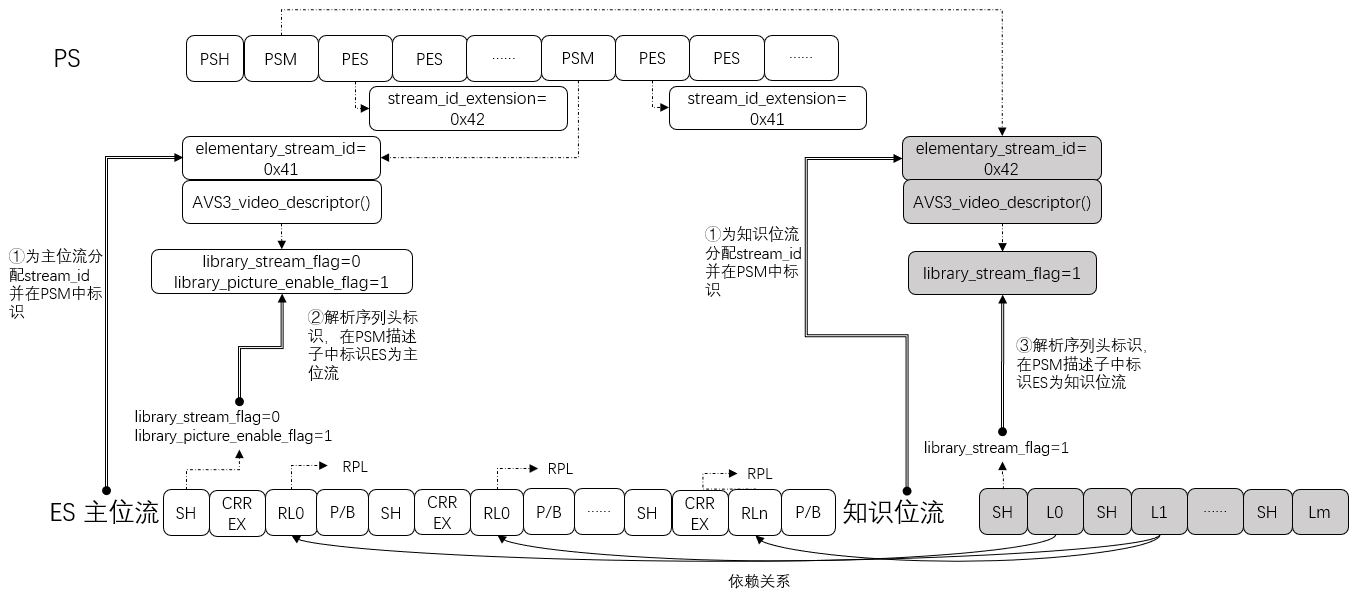
生成节目流应从视频位流中解析语法元素用于为节目流中PSM中的相关字段设定该字段应取的值。从视频位流的序列头中解析library\_stream\_flag和library\_picture\_enable\_flag的值，方法如下：

* 1. 当library\_stream\_flag的值为0且library\_picture\_enable\_flag的值为0时，判断该视频位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；
  2. 当library\_stream\_flag的值为0且library\_picture\_enable\_flag的值为1时，判断该视频位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；
  3. 当library\_stream\_flag的值为1时，判断该视频位流包含知识位流，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视频位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位流，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

* 1. 传输流解封装

AVS3视频位流的传输流解封装流程如下：

1. 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：
2. 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的PES分组包；
3. 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’；
4. 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；
5. 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：
6. 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的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的PES分组包；
7. 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。

* 1. 节目流解封装

AVS3视频位流的节目流解封装流程如下：

1. 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：
2. 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；
3. 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。
4. 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：
5. 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；
6. 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。