ICS 33.160.25
CCS M74

UHD World Association
世界超高清视频产业联盟

世 界 超 高 清 视 频 产 业 联 盟 标 准

T/UWA 038.2-2—2026

菁彩影院技术 第 2-2 部分：菁彩电影声发行母版技术要求

Vivid cinema technology part 2-2:
Technical requirements for digital cinema distribution master of audio vivid cinema

2026－11－23 发布　　2026－11－23 实施

世界超高清视频产业联盟　　发 布

目　　次

前言 II
1 范围 1
2 规范性引用文件 1
3 术语和定义 1
4 缩略语 2
5 制作系统 2
6 技术要求 3
6.1 基于声道的音频要求 3
6.2 基于声床与对象的音频要求 3
附录 A （规范性） 音频对象的空间坐标系 6
附录 B （规范性） 影院音频对象位置元数据与菁彩声位置元数据映射关系 7
参考文献 8

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。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由世界超高清视频产业联盟提出并归口。

本文件起草单位：中国电影科学技术研究所（中央宣传部电影技术质量检测所）、华夏电影发行有限责任公司、华为技术有限公司、深圳市洲明科技股份有限公司、华夏利亚德（北京）电影科技有限公司、中央宣传部电影数字节目管理中心、深圳光峰科技股份有限公司、西安诺瓦星云科技股份有限公司、深圳市奥拓电子股份有限公司、北京东方佳联影视技术有限公司。

本文件主要起草人：王薇娜、董强国、常婉晴、邢占宇、燕兴、董志刚、王加志、刘翼光、杨城、王勇、古俊豪、徐涛、龚波、肖令权、郝效荣、高鹏、李虹珊、李雪伟、周波、林先萌、刘斌、牛小明、刘昕、严振航、成相翼。

菁彩影院技术 第2-2部分：
菁彩电影声发行母版技术要求

1 范围

本文件规定了菁彩电影声发行母版的制作系统、通用音频特性制版要求。

本文件适用于菁彩电影声发行母版的制作。

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。

GB/T 46271—2025 信息技术 三维声技术 编码、分发与呈现

SMPTE ST 429—18:2019 数字电影打包 沉浸式音轨文件（D-Cinema Packaging - Immersive Audio Track File）

SMPTE ST 2098-2:2022 沉浸式音频码流规范（Immersive Audio Bitstream Specification）

3 术语和定义

下列术语和定义适用于本文件。

3.1

音频对象 audio object

一组由音频信号及其相关可随时间灵活变化元数据组成的音频内容。

注：对象的扬声器分配通常是动态的。

[来源：T/CSMPTE 35—2024，定义3.1，有修改]

3.2

声床 audio bed

组成沉浸式混音音轨基础的声场组。

注1：如7.1、9.1等。

注2：一部影片通常包含一组声床。

注3：声床的扬声器分配通常是静态的，在整个节目中不会改变。

[来源：T/CSMPTE 35—2024，定义3.2，有修改]

3.3

元数据 metadata

用于描述音频对象、声床音频的位置、增益等属性的数据。

[来源：T/CSMPTE 35—2024，定义3.3，有修改]

3.4

沉浸式音频 immersive audio

由通道、声床、对象的音频内容以及相关元数据组成，经特定渲染算法，将元数据的位置等信息在还音空间的还音效果中呈现，能构建比单一水平层面更丰富的空间音频体验。

[来源：T/CSMPTE 35—2024，定义3.4]

3.5

沉浸式音频码流 immersive audio bitstream

沉浸式音频的所有元素根据一定规则和结构编码组成的比特流。

[来源：T/CSMPTE 35—2024，定义3.5]

3.6

菁彩电影声发行母版　immersive audio for DCDM

一组用于数字电影发行包制作的菁彩电影声文件格式。

注：包含传统基于声道的音频文件、基于声床和对象的沉浸式音频码流文件。

3.7

采样率　sample rate

音频信号的采样频率。

[来源：T/CSMPTE 35—2024，定义3.7]

3.8

量化深度　bit depth

单个音频信号采样点振幅数值的有效位数。

[来源：T/CSMPTE 35—2024，定义3.8]

3.9

子块　Sub block

一个图像帧时长内，音频采样点数据集切分为5.2ms、4.2ms子集，一个子块对应一组元数据。

4　缩略语

下列缩略语适用于本文件。

DCDM：数字电影发行母版（Digital Cinema Distribution Master）

DCP：数字电影数据包（Digital Cinema Package）

DLC：动态无损编码（Dynamic Lossless Coding）

ID：标识（Identification）

MXF：素材交换格式（Material eXchange Format）

PCM：脉冲编码调制（Pulse Code Modulation）

5　制作系统

菁彩电影声制作系统示意图见图1，主要包含三部分内容：

1）声像制作工具：对音频对象的空间位置、运动轨迹、增益等参数进行精细化调控，生成对象与声床的空间声像控制元数据。元数据通过网络接口传输，声音素材通过网络声卡协议传输至下一环节。

2）混录、渲染与监听工具：接收并同步录制对象音频、声床信号与控制元数据，生成包含完整创作信息的一体化中间文件；基于元数据与音频数据，完成对象声的空间声像计算、声场渲染与混音，生成适配影院重放的多声道信号，用于制作现场监听与回放。

3）文件编码与封装工具：对制作完成的音频内容进行编码，生成符合规范的菁彩电影声编码流，将编码后的音频流封装为MXF格式文件，满足数字影院的媒体文件规范，可进一步用于DCP的制作。

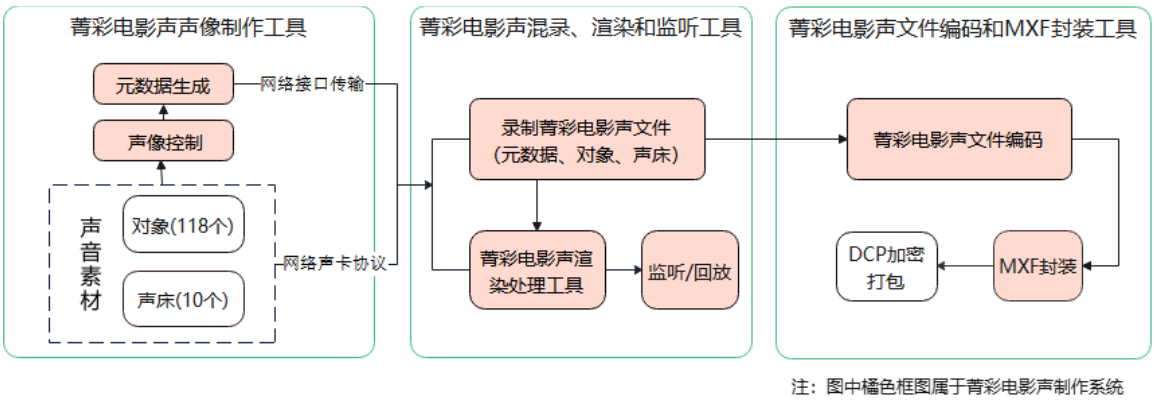


图1　菁彩电影声制作系统

6 技术要求

6.1 基于声道的音频要求

6.1.1 采样率

采样率应为48 kHz或96 kHz。

6.1.2 量化深度

每个采样点的量化深度应为24 bit。

6.1.3 数字音频参考电平

针对数字影院的母版制作，数字输入和输出的参考电平应为-20 dBFS。

注：数字音频参考电平是指统一音频系统中模数/数模转换时的基准音频电平。

6.1.4 声道音轨

基于声道的音频文件应支持16条音轨，声道配置应至少支持5.1声道或7.1声道。

5.1音频、7.1音频的声道映射应符合表1的规定。

表1 母版声道音轨映射

声道音轨顺序	5.1音频	7.1 音频
第1轨	左声道	左声道
第2轨	右声道	右声道
第3轨	中置声道	中置声道
第4轨	次低频声道	次低频声道
第5轨	左环绕声道	左侧环绕声道
第6轨	右环绕声道	右侧环绕声道
第7轨	听障辅助	听障辅助
第8轨	视障辅助	视障辅助
第9轨	—	—
第10轨	—	—
第11轨	—	左后环绕声道
第12轨	—	右后环绕声道
第13轨	—	—
第14轨	—	—
第15轨	—	—
第16轨	—	—

6.1.5 音频素材编码

音频素材应采用PCM编码，并以广播波形文件格式存储。

广播波形文件中应包含元数据以标记音频数据的第一个样本，元数据还应包含相对于图像的连续帧计数以及采样率。

6.2 基于声床与对象的音频要求

6.2.1 基础音频特性

沉浸式音频码流中，声床和对象音频的基础音频特性应符合表2的规定。

表2 声床和对象音频的基础音频特性

序号	参数	技术要求
1	采样率	应符合6.1.1的规定
2	量化深度	应符合6.1.2的规定
3	数字音频参考电平	应符合6.1.3的规定
5	音频码流帧率	应与画面帧率一致，应至少支持24fps

6.2.2 音频素材编码

基于声床和对象的音频素材编码应采用 DLC 编码，应符合 SMPTE ST 2098-2:2022 中附录 B 规定的编码规则。

注：DLC编码是一种音频素材的编码方式，该编码基于PCM编码的无损压缩格式。

6.2.3 同步信号

基于声床与对象的音频同步信号应放置于6.1.4所述声道音轨的第14轨中。

6.2.4 声床和对象音轨数量

基于声床和对象的音频应支持128条音频素材音轨。

6.2.5 声床制式与元数据

声床的声道制式应至少支持7.1声床制式或9.1声床制式，声床元数据的通道ID应符合表3的规定。

表3 声床制式与声床元数据通道 ID

序号	声道名称	通道ID	7.1	9.1	7.1.4
1	左声道	0x0	√	√	√
2	右声道	0x4	√	√	√
3	中置声道	0x2	√	√	√
4	次低频声道	0xD	√	√	√
5	左环绕声道	0x5	√	√	√
6	右环绕声道	0x9	√	√	√
7	左后环绕声道	0x7	√	√	√
8	右后环绕声道	0x8	√	√	√
9	左顶环绕声道	0xB	—	√	—
10	右顶环绕声道	0xC	—	√	—
11	左前顶环绕声道	0x18			√
12	右前顶环绕声道	0x19			√
13	左后顶环绕声道	0x1A			√
14	右后顶环绕声道	0x1B			√
注：“√”表示该声床中包含此声道。					

6.2.6 对象元数据

沉浸式音频对象元数据描述的特性应符合表4的规定。

表4 对象元数据描述特性

序号	特性	描述	技术要求	
			影院用	非影院用
1	空间坐标	用于描述音频对象空间位置的坐标系	应符合附录A的规定	
2	扩散特性	用于表示音频对象在空间中大小与形状随时间变化的特性	应至少支持1维模式的扩散特性	可支持
3	去相关特性	用于表示音频信号去相关的程度，去相关处理可以使多个扬声器播放多个关联的音频内容时，产生更宽、更分散的声像	应支持	可支持
4	区域特性	用于将还音空间划分区域，按区域分配音频增益	可支持	
5	对象的声道吸附特性	用于描述音频对象对还音扬声器位置的限定范围	可支持	
注：“影院用”指在符合电影行业规范的商业电影院放映环境中使用；“非影院用”指影院以外的一切使用场景，包括但不限于家庭影院等近场混音场景中使用。				

6.2.7 子块

不同采样率与图像帧速率条件下，子块数量与子块中的采样点数应符合表5的规定。

表5　不同采样率、图像帧速率对应的子块数量与采样点数

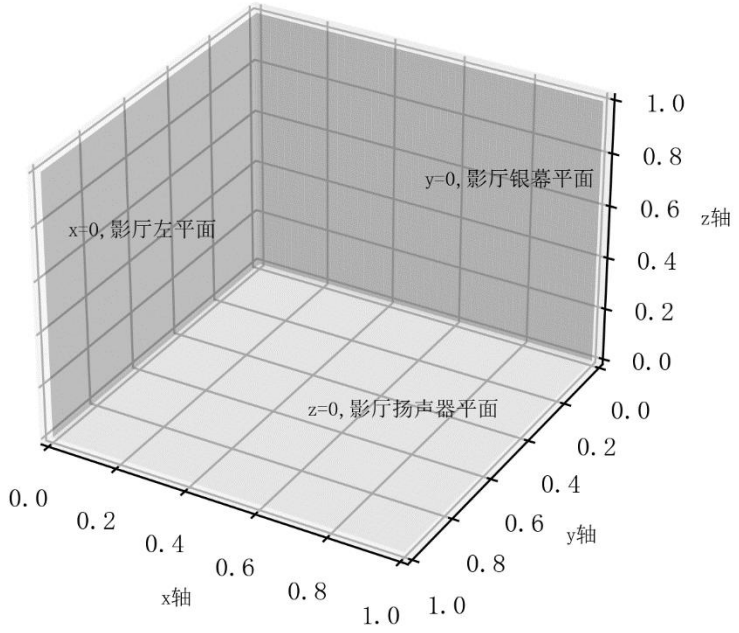
序号	采样频率	图像帧速率	子块时长	子块中的采样点数	子块数
1	48 kHz	24	5.2 ms	250	8
2		48	5.2 ms	250	4
3		60	4.2 ms	200	4
4	96 kHz	24	5.2 ms	500	8
5		48	5.2 ms	500	4
6		60	4.2 ms	400	4

6.2.8　码流编码

菁彩电影声音频基于声床和对象的码流编码，应符合SMPTE ST 2098-2:2022的规定，编码后封装于MXF文件中，封装要求应符合SMPTE ST 429-18:2019的规定。

附 录 A
（规范性）
音频对象的空间坐标系

菁彩电影声音频对象应使用笛卡尔坐标系表示音频对象的位置，该坐标系使用三个正交轴（x，y，z）来定位空间中相对于原点的一个点。其中，x轴代表影厅横向/左右位置，y轴代表影厅纵向/前后位置，z轴代表影厅高度/上下位置，见图A.1。



图A.1 对象空间坐标系（左手坐标系）

（x，y，z）坐标值范围从（0,0,0）到（1,1,1），原点为影厅的左前角，典型坐标位置与影厅空间的对应关系应符合表A.1中的规定。

表A.1 典型坐标位置与影厅空间的对应关系

典型坐标值	空间对应位置
x=0	影厅左侧墙壁
x=1	影厅右侧墙壁
y=0	影厅银幕侧墙壁
y=1	影厅后侧墙壁
z=0	影厅中5.1、7.1系统扬声器所在高度中心平面
z=1	影厅天花板

附 录 B
（规范性）
影院音频对象位置元数据与菁彩声位置元数据映射关系

GB/T 46271-2025规定了菁彩声的编解码及渲染标准，本标准元数据应符合GB/T 46271-2025中元数据定义。

菁彩声音频对象位置坐标系的原点为甜点位置，即测听者位于房间中心位置，且面向屏幕方向。（x，y）平面位于主声道和环绕声道扬声器系统声中心所在水平面。（x，y，z）坐标轴中，x轴为左右方向，其中右方为正方向，y轴为前后方向，其中前方为正方向，z轴为上下方向，其中上方为正方向。x、y、z轴坐标的取值范围均为[-1,1]。

将附录A中影院音频对象位置元数据转换为菁彩声位置元数据时，转换公式见公式（B.1）～公式（B.3）。

$$x_{vivid} = 2.0 \times x_{th} - 1.0 \quad \text{(B.1)}$$

式中：

x_{vivid}——菁彩声坐标系下的对象x坐标；

x_{th}——影院音频对象的x坐标。

$$y_{vivid} = 2.0 \times y_{th} + 1.0 \quad \text{(B.2)}$$

式中：

y_{vivid}——菁彩声坐标系下的对象y坐标；

y_{th}——影院音频对象的y坐标。

$$z_{vivid} = z_{th} \quad \text{(B.3)}$$

式中：

z_{vivid}——菁彩声坐标系下的对象z坐标；

z_{th}——影院音频对象的z坐标。

将菁彩声对象位置元数据转换为影院对象位置元数据时，转换公式见公式（B.4）～公式（B.6）。

$$x_{th} = 0.5 \times x_{vivid} + 0.5 \quad \text{(B.4)}$$

$$y_{th} = -0.5 \times y_{vivid} + 0.5 \quad \text{(B.5)}$$

$$z_{th} = \begin{cases} z_{vivid} & if\ z_{vivid} \geq 0 \\ 0 & if\ z_{vivid} < 0 \end{cases} \quad \text{(B.6)}$$

其中，由于菁彩声对象位置元数据z轴的取值范围为[-1,1]，影院位置元数据z轴的取值范围为[0,1]，因此z_{vivid}转换为z_{th}时，需要进行空间位置坐标体系的转换。

参 考 文 献

[1] GY/T 292.2—2015 数字电影发行母版 第2部分：音频特性

[2] GY/T 312—2017 电影 录音控制室、室内影厅B环电声响应规范和测量

[3] T/CSMPTE 35—2024 数字电影沉浸式音频发行母版技术要求

[4] SMPTE ST 2098-5:2018 数字电影沉浸式音频通道和声场组（D-Cinema Immersive Audio Channels and Soundfield Groups）

[5] SMPTE RDD 52:2020 D-Cinema Packaging — DCP Bv2.1 Application Profile[R].SMPTE, 2020.

[6] SMPTE RDD 57:2021 SMPTE ST 2098-2 Immersive Audio Bitstream and Packaging Constraints: IAB Application Profile 1[R].SMPTE, 2021.