

ICS 号: xxxx-xxxx  
中国标准文献分类号



# 世界超高清视频产业联盟标准

T/UWA xxxx-2024

## 三维声音技术规范 第 3-x 部分:

### 技术要求与测试方法 菁彩声混音棚

3D Audio Technology Specification Part 3-x:

Technical Requirement and Test Method – Audio Vivid Mixing Studio

(征求意见稿)

2024-xx-xx 发布

2024-xx-xx 实施

世界超高清视频产业联盟

# 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
3.1 三维声 3D Audio .....	1
3.2 三维菁彩声 Audio Vivid .....	1
4 缩略语 .....	1
5 概述 .....	2
6 技术要求 .....	2
6.1 听音环境要求 .....	2
6.2 混音系统软硬件要求 .....	2
6.3 视音频相对延时控制 .....	4
7 一般测试条件 .....	4
7.1 环境条件 .....	4
7.2 电源 .....	4
7.3 额定工作状态的调整 .....	4
7.4 测量接口 .....	4
7.5 稳定时间 .....	4
7.6 主要测量仪器 .....	4
8 测试信号 .....	5
8.1 听音环境测试信号 .....	5
8.2 输入文件支持相关测试信号 .....	5
8.3 视音频相对延时测试信号 .....	6
9 测量方法 .....	6
9.1 听音环境测量方法 .....	6
9.2 混音系统软硬件测量方法 .....	6
9.3 视音频相对延时控制测量方法 .....	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》给出的规则起草。

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这些专利的责任。

本文件由世界超高清视频产业联盟提出并归口。

本文件主要起草单位：国家广播电视总局广播电视规划院、中央广播电视总台、中国电子技术标准化研究院、华为技术有限公司、赛因芯微（北京）电子科技有限公司、音王电声股份有限公司、世界超高清视频产业联盟、中广电融合（北京）科技有限公司、央广云听文化传媒有限公司、上海数字电视国家工程研究中心、北京流金岁月传媒科技股份有限公司、深圳创维-RGB 电子有限公司、咪咕文化科技有限公司、杭州当虹科技股份有限公司、北京数码视讯科技股份有限公司、视源电子科技有限公司、中国美术学院电影学院、上海交通大学文创学院、中国传媒大学音乐与录音艺术学院、鼎盛佳和（北京）文化传播有限公司、北京一灵二聆科技发展有限公司、浙江传媒学院电视艺术学院、迈思非墨（北京）文化传媒有限公司

本文件主要起草人：

# 三维声音技术规范 第 3-x 部分： 技术要求与测试方法 菁彩声混音棚

## 1 范围

本文件规定了三维菁彩声（Audio Vivid）混音用混音棚的听音环境、混音系统软硬件等技术要求和相关测试方法。

本文件适用于三维菁彩声（Audio Vivid）音频内容的制作，也适用于三维菁彩声（Audio Vivid）制作系统搭建及设备的设计、生产、验收、运行和维护。

## 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 13581-1992 高保真头戴耳机最低性能要求

GB/T 14471-2013 头戴耳机通用规范

GB/T 42758-2023 用于节目制作的先进声音系统

GB/T 19889.4-2005 声学 建筑和建筑构件隔声测量 第 4 部分：房间之间空气声隔声的现场测量

GY J42-89 声学背景噪声要求

GY 5022-2007 广播电视播音（演播）室混响时间测量规范

GY/T 298-2016 音频系统小损伤主观评价方法

5086GY/T 363-2023 三维声编解码及渲染

GY/T 5086-2012 广播电视录（播）音室、演播室声学设计规范

ITU-R BS.2076 声音定义模型（Audio Definition Model）

## 3 术语和定义

下列术语以及定义适用于本文件。

### 3.1 三维声 3D Audio

一种音频格式，该格式下，多个声道承载构成完整音频内容的多路音频信号，通过环绕听音者的位于不同高度层的多个扬声器直接重放，或经过渲染或映射后重放，提供更高的声像定位空间解析度，并给听音者带来沉浸式的声场感觉。

### 3.2 三维菁彩声 Audio Vivid

GY/T 363-2023 规定的音频编解码技术规范，及配套衍生技术的代称。

## 4 缩略语

下列缩略语适用于本文件。

DAW 数字音频工作站（Digital Audio Workstation）

HOA 高阶立体声场（Higher Order Ambisonics）

ADM 音频设备管理模块（Audio Device Manager）

## 5 概述

混音棚内典型三维菁彩声（Audio Vivid）音频制作系统构成如图 1 所示。声音信号（包括音频文件、音频实时信号和 Audio Vivid 解码器解码后文件等）输入数字音频工作站，在数字音频工作站中进行声道分配、声像定位等处理后，将三维菁彩声信号（包含声床、对象和元数据信息）发送到三维菁彩声渲染器进行声音和元数据的渲染（通过扬声器或耳机进行渲染监听），也可通过 Audio Vivid 编码器压缩编码后输出。

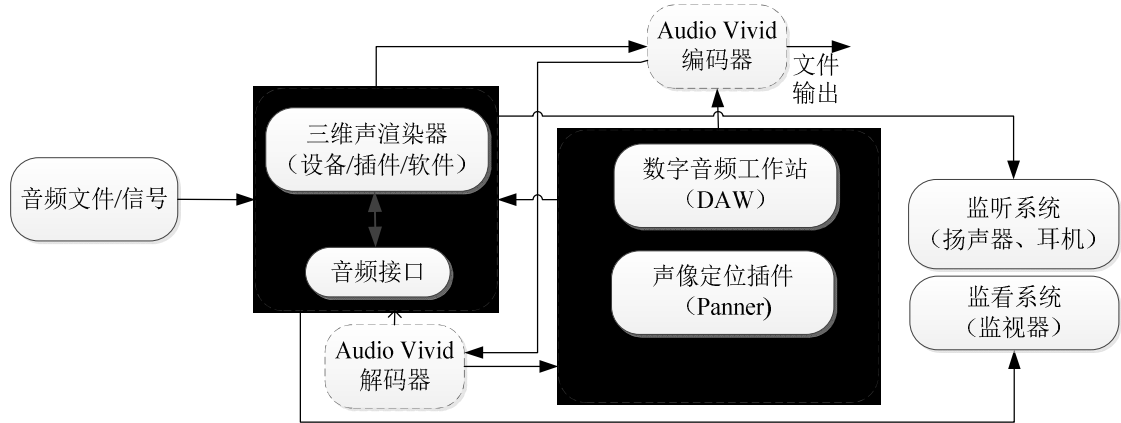


图 1 混音棚内典型三维菁彩声（Audio Vivid）音频制作系统构成

## 6 技术要求

### 6.1 听音环境要求

#### 6.1.1 房间条件

设置监听音箱时，混音棚的面积宜为 30~70m<sup>2</sup>，应大于 10 m<sup>2</sup>。房间高度宜≥2.6 米。房间比例应符合 T/CSMPTE 5—2018 表 1 的要求。

#### 6.1.2 声学特性

房间的混响时间应符合 GY/T 5086-2012 中“4.1.2 文艺类录音室混响时间”的规定。背景噪声宜满足 NR20，至少应满足 NR25。计权空气声隔声量  $D_{nT,W}$  不宜低于 60dB，至少应满足 55dB。

#### 6.1.3 扬声器配置和布局

扬声器配置和布局应符合 GB/T 42758-2023 的要求，扬声器布局宜为 5.1.4。扬声器电声性能宜符合 GY/T 298-2016 9.2.2 的要求。

#### 6.1.4 监听耳机要求

监听耳机与重放质量直接有关特性的相关要求应符合 GB/T 13581-1992 和 GB/T 14471-2013 的要求。

### 6.2 混音系统软硬件要求

#### 6.2.1 输入文件支持和输出文件要求

混音系统应支持多种声音格式的输入，可包括对象、多声道声床以及HOA信号等，并支持元数据解析，至少应包括GY/T 363-2023、ITU-R BS.2076规定的元数据；支持三维菁彩声（Audio Vivid）编码，应支持基于声道的编码、基于对象的编码和ADM元数据，宜支持HOA空间编码；应支持扬声器渲染和双耳渲染；应支持导出WAV、ADM及AV3A（Audio Vivid）格式文件。具体功能应符合表1要求。

表1 混音棚输入输出要求

序号	项目	技术要求
1	音频格式支持	应支持三维声（基于声床、基于对象），以及环绕声、立体声、单声道
2	编解码和渲染	应支持三维菁彩声（Audio Vivid）有损和无损编解码
		应支持扬声器渲染和双耳渲染
		宜支持3/6自由度（3DoF/6DoF）呈现
		应支持ADM元数据
		宜支持三维菁彩声（Audio Vivid）扩展元数据
		宜支持HOA空间编码工具（HOA可支持到3阶）
		应支持至少16个通道的编解码
		编解码码率应支持32kbps~1.6Mbps
3	采样率	应支持32kHz~192kHz
4	位宽	应≥16比特

## 6.2.2 硬件功能要求

### 6.2.2.1 音频接口功能

音频接口应至少支持1组10通道的输出。

### 6.2.2.2 监听控制器功能

监听控制器应至少支持1组10（5.1.4）通道的输出。应具备如表2所示功能。

表2 监听控制器功能

序号	项目	功能要求
1	音量控制	应具备音量控制功能，并对音量进行量化显示
		宜对独立通道进行独奏/静音（Solo/Mute）切换
2	信号源切换	应具备信号源切换功能
3	其他功能	应具备均衡调整选项
		宜支持对电平进行测量并显示

## 6.2.3 软件功能要求

混音棚音频系统的软件应与三维菁彩声（Audio Vivid）的混录功能相匹配，具体应具备功能如表3所示。

表3 软件功能要求

序号	项目	功能要求
1	三维空间声像	应具备三维空间声道定位功能
		应具备三维空间声像调整功能
2	音频对象的输出	应具备音频对象和元数据的渲染输出功能
		应具备音频对象和ADM元数据的导出功能
3	声床的输出	应具备声床和元数据的渲染输出功能
		应具备声床和ADM元数据的导出功能
4	三维菁彩声（Audio Vivid）的输出	应具备16通道三维菁彩声（Audio Vivid）的自定义导出功能
		应具备指定音频对象可交互元数据的设置和导出功能
		应具备Audio Vivid编码导出功能
5	双耳渲染输出	应具备声床、音频对象和ADM元数据的导出功能
6	渲染监听	应具备多扬声器与双耳渲染的实时监听功能

### 6.3 视音频相对延时控制

应保证监听监看系统的音视频同步，视音频相对延时应在-30ms~22.5ms范围内（负号表示音频滞后）。

## 7 一般测试条件

### 7.1 环境条件

应在下列温度、湿度和气压条件范围内进行测试：

- 环境温度：15~35℃
- 相对湿度：25~75%RH
- 气压：86~106kPa

### 7.2 电源

测量应在额定电源电压条件下，测试时电源电压的变化不超出±2%；当采用交流电网供电时，电源频率的波动不超出±2%，谐波分量不超出5%。

### 7.3 额定工作状态的调整

听音环境测试时，混音棚内空调、通风、灯光等产生噪声的设备关闭。声学分析仪应放置在被测混音棚之外，避免测试仪器本身发出的噪声影响测量结果，同时被测混音棚内应避免人为的噪声影响测量结果。

### 7.4 测量接口

采用文件形式输入多声道音频至被测混音棚，优选USB。

### 7.5 稳定时间

在测试前，应使被测试设备在额定测量条件下工作15min，以使其性能稳定。

### 7.6 主要测量仪器

#### 7.6.1 音频分析仪

音频分析仪应具备以下功能：

- a) 具备600Ω输入阻抗选项，22kHz或30kHz低通滤波器，997Hz 1/3oct带通滤波器，A计权滤波器选件；
- b) 具备电压表、相位表和失真仪功能；
- c) 具备FFT频谱分析功能；
- d) 具备波形监测功能。
- e) 具备数字音频输入功能。

注1：如不具备FFT频谱分析功能，可用20Hz~20kHz频段的频谱分析仪。

注2：如不具备波形监测功能，可用示波器。

#### 7.6.2 声压级测量设备

声压级测量设备应具备以下功能：

- a) 设备精度应符合GB/T 3785-1983和GB/T 17181-1997中规定的0型或1型的准确度要求，这个测量系统在每次测量前使用符合GB/T 15173-1994规定的1级精度要求的校准器进行校准。

- b) 滤波器应符合GB/T 3241-1998的要求。
- c) 混响时间的测量设备应符合ISO 354:1985规定的要求。

## 8 测试信号

### 8.1 听音环境测试信号

听音环境测试信号应符合表3的规定。

表4 听音环境测试信号

序号	测试信号	编码	封装	信号描述
1	Speaker_crossover_m.0.n	Audio Vivid	.TS	m加n个全带宽声道，各声道信号在时间轴上依次出现粉红噪声

注：其中m表示耳部高度层（或稍低）全带宽声道的数量，n表示顶部全带宽声道的数量。

### 8.2 输入文件支持相关测试信号

#### 8.2.1 声道数支持测试信号

声道数支持测试信号应符合表4的规定。

表5 声道数支持测试信号

序号	测试信号	编码模式	声道设置
1	Channel_ID_Voice_2.0	2.0	L,R
2	Channel_ID_Voice_5.1	5.1	L,R,C,LFE,Lss,Rss
3	Channel_ID_Voice_5.1.4	5.1.4	L,R,C,LFE,Ls,Rs,Ltf,Rtf,Ltb,Rtb
4	Channel_ID_Voice_7.1.4	7.1.4	L,R,C,LFE,Lss,Rss,Lrs,Rrs,Ltf,Rtf,Ltb,Rtb

#### 8.2.2 声音信号输入测试信号

声音信号输入测试信号应符合表5的规定。

表6 声音信号输入测试信号

序号	测试信号		编码模式	音频编码
1	Object_4_5.1.4	声床	5.1.4	Audio Vivid
2		对象	对象 <sup>注</sup>	Audio Vivid
3	Object_4_7.1.4	声床	7.1.4	Audio Vivid
4		对象	对象	Audio Vivid
5	HOA_3order		HOA（3阶）	Audio Vivid

注：对象应包含静态对象和动态对象。其中静态对象ADM中包含的位置信息应对应固定扬声器位置，动态对象ADM中包含的位置信息应随时间不断变化。

#### 8.2.3 信源格式支持测试信号

##### 8.2.3.1 采样率和量化比特

采样率和量化比特测试信号应符合表6的规定。

表7 采样率和量化比特测试信号

序号	测试信号	编码	封装	信号描述
1	Source_SR_32kHz_16bit	LPCM	.WAV	2声道，32kHz采样率，16bit量化
2	Source_SR_32kHz_24bit			2声道，32kHz采样率，24bit量化
3	Source_SR_44.1kHz_16bit			2声道，44.1kHz采样率，16bit量化
4	Source_SR_44.1kHz_24bit			2声道，44.1kHz采样率，24bit量化
5	Source_SR_48kHz_16bit			2声道，48kHz采样率，16bit量化



6	Source_SR_48kHz_24bit			2声道, 48kHz采样率, 24bit量化
---	-----------------------	--	--	------------------------

### 8.2.3.2 Audio Vivid 菁彩声比特率

Audio Vivid菁彩声比特率支持测试信号应符合表7的规定。

表8 Audio Vivid菁彩声比特率支持测试信号

序号	测试信号	编码	封装	信号描述	备注
1	Vivid_bps_2.0_48	Audio Vivid	.TS	2声道, 48kbps	基本流为48kHz, 24bit
2	Vivid_bps_2.0_320			2声道, 320kbps	
3	Vivid_bps_5.1_256			5.1声道, 256kbps	
4	Vivid_bps_5.1_720			5.1声道, 720kbps	
5	Vivid_bps_7.1_384			7.1声道, 384kbps	
6	Vivid_bps_7.1_640			7.1声道, 640kbps	
7	Vivid_bps_7.1.4_384			7.1.4声道, 384kbps	
8	Vivid_bps_7.1.4_704			7.1.4声道, 832kbps	

### 8.3 视音频相对延时测试信号

视音频相对延时测试信号应符合表8的规定。

表9 视音频相对延时测试信号

序号	测试信号	编码	封装	信号描述
1	avdelay_48k_16bit	Audio Vivid	.ts	视音频相对延时量为0的测试信号

## 9 测量方法

### 9.1 听音环境测量方法

房间的混响时间应采用GY 5022-2007的测量方法。

背景噪声应采用GY J42-89的测量方法。

两室间的空气声隔声量应采用GB/T 19889.4-2005的测量方法。

### 9.2 混音系统软硬件测量方法

#### 9.2.1 测量框图

输入输出信号及文件支持测量框图如图2所示。

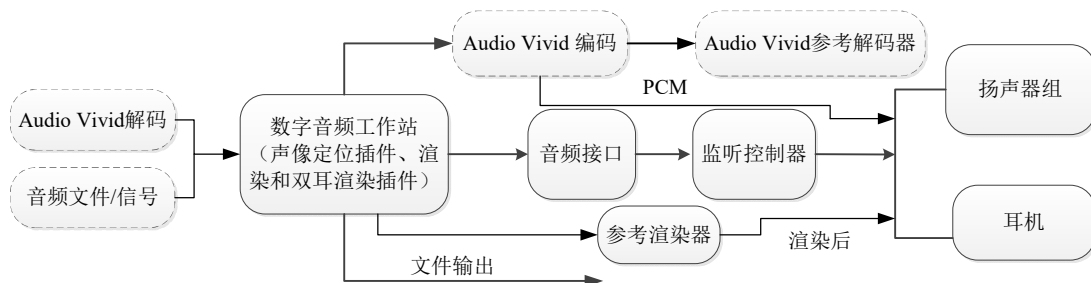


图2 功能验证测量框图

#### 9.2.2 输入文件支持和输出文件支持测量步骤

输入输出信号及文件支持测量步骤如下：

- a) 按照图2将8.2.2的测试信号输入至数字音频工作站，导出含ADM的PCM音频，经参考渲染器渲染后监听。
- b) 按照图2将8.2.2的测试信号输入至数字音频工作站，经三维菁彩声（Audio Vivid）编码后输出码流，由参考解码器解码，验证渲染监听是否正常。

### 9.2.3 硬件功能测量方法

#### 9.2.3.1 音频接口功能测量方法

按照图2在混棚内连通音频接口，由数字音频工作站通过音频接口将声音信号输出至扬声器组或耳机，验证是否支持1组10通道的输出。

#### 9.2.3.2 监听控制器功能测量方法

监听控制器功能的测量步骤如下：

- a) 按照图2将8.2.2的测试信号输入至数字音频工作站，在监听控制器端验证其是否支持1组10路信号的对比监听；
- b) 监听过程中调整监听控制器音量控制选项，验证其是否正确控制音量；
- c) 监听过程中对独立通道进行独奏/静音（Solo/Mute）通道切换，验证是否切换正常；
- d) 验证监听控制器是否具备均衡调整功能及是否支持电平表和响度表。

### 9.2.4 软件功能测量方法

测量步骤如下：

- a) 验证数字音频工作站是否支持三维空间的声道定位及声像调整。
- b) 验证数字音频工作站能否导出音频对象和ADM元数据。
- c) 验证是否支持对象和元数据的渲染输出及双耳渲染输出。

## 9.3 视音频相对延时控制测量方法

### 9.3.1 测量框图

测量框图见图3。

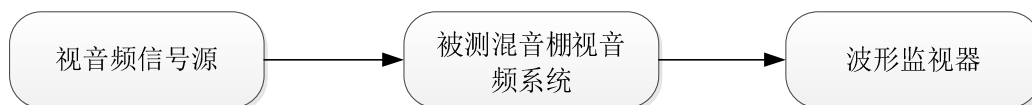


图3 音视频相对延时测量框图

### 9.3.2 测量步骤

测量步骤如下：

- a) 按图3连接测量仪器和被测系统；
- b) 视音频信号源输出视音频相对延时量为零的测量信号；
- c) 视音频相对延时测量信号经被测系统后，由波形监视器测量出音视频相对延时。

## 参 考 文 献

- [1] GB/T 50121-2005 建筑隔声评价标准
- [2] GY/T 316-2018 用于节目制作的先进声音系统
- [3] ITU-R BS.1909-0 对带有或不带伴图的先进多声道立体声系统的性能要求  
(Performance requirements for an advanced multichannel stereophonic sound system for use with or without accompanying picture)
- [4] ITU-R BS.2051 用于节目制作的先进音响系统 (Advanced sound system for programme production)
- [5] 中国电子音响行业协会高保真音频评测评价规范
- [6] T/UWA 009.1-2022 三维声音技术规范 第1部分：编码分发与呈现
- [7] T/UWA 009.3-1-2022 三维声音技术规范 第3-1部分：技术要求和测试方法 家庭影音播放设备
- [8] T/CSMPTE 5-2018 基于4K超高清图像和环绕声/三维声的家庭影院配置规范