

世界超高清视频产业联盟团体标准

T/UWA XXXX—XXXX

车载空间音频有声剧制作标准

Production Standard for In-Vehicle Spatial Audio Drama

（征求意见稿）

XXXX – XX – XX 发布

XXXX – XX – XX 实施

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..... 1

4 缩略语..... 1

5 制作环境..... 1

5.1 硬件环境..... 1

5.2 软件环境..... 2

5.2.1 核心功能要求..... 2

5.2.2 格式兼容性 2

6 工程文件制作规格与要求..... 2

6.1 工程制作音源要求..... 2

6.2 声床 / 对象 要求..... 2

6.3 低频通道要求..... 2

6.4 声像定位要求..... 2

6.5 车载适配要求..... 3

7 分发编码与交付..... 3

7.1 分发编码要求..... 3

7.2 交付要求..... 3

7.2.1 交付文件格式 3

7.2.2 分发文件响度要求 3

附 录 A 空间音频有声剧制作流程..... 4

A.1 制作流程参考..... 4

附 录 B 人声/音效声场均匀度优化测试方法..... 5

B.1 测试目的..... 5

B.2 测试准备..... 5

B.3 测试信号..... 5

B.4 测量方法..... 5

B.5 数据处理要求..... 6

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。

本文件由世界超高清视频产业联盟提出并归口。

本文件起草单位：上海喜马拉雅科技有限公司、中央广播电视总台、国家广播电视总局广播电视规划院、华为技术有限公司、中汽智能科技（天津）有限公司、长城汽车股份有限公司、吉利汽车控股有限公司、奇瑞汽车股份有限公司、北京流金岁月传媒科技股份有限公司、声响节拍科技有限公司、马栏山音视频实验室、上海大宁资产经营(集团)有限公司

本文件主要起草人：

引 言

随着汽车产业与音频技术的深度融合，车载场景已成为音频内容消费的重要阵地，其中兼具叙事性与听觉沉浸感的空间音频有声剧，正逐渐成为提升用户乘车体验的核心内容形态之一。然而，M1/M2 类车辆（以个人 / 家庭出行为主的轻型载客汽车）的车内声学环境具有 “空间紧凑、座位布局集中、行驶中存在环境噪声干扰” 等独特属性，且驾驶场景对 “不分散注意力、不干扰路况判断” 存在硬性安全约束，这使得车载空间音频有声剧的制作面临特殊挑战。

当前行业内，车载音频内容的制作标准主要聚焦于音乐（侧重氛围渲染）与有声书（侧重单一信息传递），而有声剧作为 “含旁白、对白、音乐、音效的多元素协同叙事载体”，既无法直接套用车载音乐的 “氛围化制作逻辑”（易因音效过度动态导致注意力分散），也不能沿用有声书的 “单一信息传递逻辑”（难以体现空间音频的三维沉浸优势）。缺乏专属技术标准的现状，已导致市场上的车载有声剧出现 “声场不均匀（后排乘客听不清剧情）、关键声像定位偏差（干扰驾驶判断）” 等问题，不仅影响用户体验，更潜藏驾驶安全风险，同时也增加了制作方与车企的技术对接成本，制约了车载有声剧产业的规模化发展。

基于此，本标准旨在制定 M1/M2 类车辆车载空间音频有声剧从前期准备（环境、人员、音源）、核心制作（声床配置、声效定位、低频控制）到编码交付的全流程技术要求，通过明确技术边界与适配规范，填补行业空白。

本标准的制定可实现三大核心价值：

1. 统一行业制作口径，保障体验一致性：通过规范声场均匀度、声像定位范围等关键指标，避免因无专属标准导致的有声剧车载体验参差不齐，解决 “部分座位听感差、剧情声效干扰驾驶” 等痛点，为用户提供稳定、优质的跨车型听觉体验；
2. 平衡沉浸与安全，守住驾驶场景底线：针对车载场景的安全约束，明确 “重点声像（如旁白）定位在前方 $\pm 30^{\circ}$ 范围” 等要求，在保留有声剧 “多角色演绎、三维声场沉浸感” 的同时，确保不分散驾驶员注意力、不干扰路况判断，解决 “沉浸感与安全性” 的核心矛盾；
3. 降低对接成本，推动产业规模化落地：清晰界定 M1/M2 类车辆适配的技术参数，为制作方提供明确的技术指引，减少与车企的对接试错成本，同时为车载音频系统设计提供内容适配依据，加速车载空间音频有声剧的商业化应用与产业生态构建。

车载空间音频有声剧制作标准

1 范围

本文件规定了 M1/M2类车辆（定义见 3.1）车载空间音频有声剧的制作全流程要求，覆盖从制作前期准备、核心制作到编码交付的全链路。

本文件适用于 M1/M2 类车辆车载空间音频有声剧的生产、测试与分发。

本文件可作为 M1/M2 类车辆车载音频系统设计、有声剧制作方技术对接的依据，供音频平台、内容制作机构共同遵循。

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 14471-2013 头戴耳机通用规范

GB/T 15089-2001 机动车辆及挂车分类

GY/T 377-2023 网络视听节目音频响度技术要求和测量方法

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）

T/UWA 009.3-4-2024 三维声音技术规范 第 3-4 部分：技术要求和测试方法 车载音频系统

T/UWA 009.3-5-2024 三维声音技术规范 第 3-5 部分：技术要求和测试方法 菁彩声混音棚

3 术语和定义

下列术语和定义适用于本文件。

3.1 M1/M2 类车辆 M1/M2 Category Vehicle

国标GB/T 15089-2001《机动车辆及挂车分类》规定的M1/M2类车辆。

注：M1类：座位≤9座且总质量≤3.5吨 M2类：座位>9座且总质量≤5吨。

3.2 车载空间音频有声剧 In-Vehicle Spatial Audio Drama

以车载场景为播放载体，包含人声（旁白、对白）、音乐、音效等多元素材，采用空间音频技术制作的有声剧内容。

4 缩略语

下列缩略语适用于本文件。

ADM 声音定义模型（Audio Definition Model）

LFE 低频效果（Low-Frequency Effects）

LRA 响度动态范围（Loudness Range）

LU 响度单位（Loudness Unit）

5 制作环境

5.1 硬件环境

混音棚基础要求应符合 T/UWA 009.3-5-2024 规范，房间面积应 $\geq 10\text{m}^2$ （推荐 $30\sim 70\text{m}^2$ ）、高度应 $\geq 2.6\text{m}$ ，背景噪声应 $\leq \text{NR25}$ （优先满足 NR20），混响时间应符合 GY/T 5086-2012 中录音室要求，避免原始素材受环境噪声或反射声干扰。

监听系统适配要求如下：

- a) 扬声器：应符合 T/UWA 009.3-5-2024 推荐的 7.1.4 声道配置；
- b) 耳机：应符合 GB/T 14471-2013 要求。

5.2 软件环境

5.2.1 核心功能要求

软件应具备以下核心能力：

- a) 应具备空间音频的实时渲染能力；
- b) 应支持声床和声音对象的声像定位调整功能，可对音频元素的空间位置进行精准调控；
- c) 应提供 ADM BWF .wav 文件文件输出功能，满足三维声场景信息的录入与修改需求。

5.2.2 格式兼容性

应兼容 WAV、FLAC 等主流音频格式输入。

6 工程文件制作规格与要求

6.1 工程制作音源要求

- a) 素材格式：宜使用高保真原始素材，宜采用格式为 48khz 采样率，24bit 量化比特的 WAV、FLAC；
- b) 多声道适配：使用多声道原始素材时，应确保声道映射与制作环境的监听系统匹配
- c) Ambisonics 适配：使用 Ambisonics 原始素材时，应确保声道映射与制作环境的监听系统匹配

6.2 声床 / 对象 要求

- a) 声床配置：空间音频有声剧工程文件宜采用 7.1.2 及以上声床，基础声道布局应符合 Audio Vivid 或其他对应空间音频技术规范，确保声场覆盖车内不同座位区域；
- b) 声床分配要求：不应通过简单声道复制将单声道素材拓展为多声道，防止相位失真影响空间特性；
- c) 声道数限制：声床声道数与音频对象声道数总和不应超出空间音频制作软件的解码与渲染能力；
- d) 对象属性：音频对象可包含静态与动态两类，应标注位置变化轨迹元数据。

6.3 低频通道要求

低频效果（LFE）通道应仅限 120Hz 以下频段内容，全频内容（如人声、乐器基音等）不应分配至 LFE 通道。

6.4 声像定位要求

- a) 稳定性验证：应在符合 5.1.1 中要求的混音棚中扬声器渲染、双耳渲染两种模式下验证声像稳定性，保证无声像抖动等问题；

- b) 重点声像定位：应在符合 5.1 中要求的混音棚中验证重点声像（如旁白）的声像定位在大部分时间维持在混音位置正前方的 $\pm 30^\circ$ 范围内。除重点声像定位要求外，动态音效的移动范围避免跨越驾驶位正前方核心区域，减少视觉分心关联影响。
- c) 音效定位要求：动态音效的移动宜在大部分时间维持在混音位置正前方的 $\pm 90^\circ$ 核心区域，减少对行车的干扰。

6.5 车载适配要求

- a) 人声及音效声场均匀度优化：针对不同座位（驾驶位、副驾位、后排中间座位）的声场，应调整制作时各声道音量平衡，具体测量方法和推荐范围参考附录 B；
- b) 多车型人声及音效声场均匀度优化：宜至少在 M1/M2 车型各一种车型上分别进行人声及音效声场均匀度检查；
- c) 宜使用“车载模拟双耳渲染”对支持但不具备测试回放条件的车型进行声场模拟，确保对应车型的驾驶位、后排位等不同乘车视角的听觉效果。测量目的和标准与附录 B 保持一致。

7 分发编码与交付

7.1 分发编码要求

- a) 参数范围：分发编码后音频的采样率宜为 48kHz；
- b) 编码格式：应确保与车载音频系统的兼容性。以 Audio Vivid（菁彩声）编码为例，应选择对应编码格式 AV3A；
- c) 编码通道数：分发编码后的声床声道数与音频对象声道数总和不应超出其相对应的解码器和渲染器的解码与渲染能力。以 Audio Vivid（菁彩声）编码为例，分发编码后的声床声道数与音频对象声道数总和不应超出 16 个。

7.2 交付要求

7.2.1 交付文件格式

- a) 母版文件：宜提交并保留 ADM BWF .wav 文件，ADM BWF .wav 文件应包含完整元数据（含声床 / 对象声道配置、动态声像轨迹）；
- b) 分发文件：应根据车载系统算力，转码为对应配置的 AV3A 格式。

7.2.2 分发文件响度要求

- a) 基于 GY/T 377-2023《网络视听节目音频响度技术要求和测量方法》中规定，交付的空间音频编码后文件的平均响度应保持一致，应为 -15LKFS 或 -24LKFS，容差范围应不超过 $\pm 2LU$ 。网络平台也可同时提供和分发 -15LKFS 和 -24LKFS 两种版本供用户选择接收嘈杂环境适配标准；编码后的文件平均响度值宜以元数据的形式写入分发文件。
- b) 音频电平真峰值应不超过 -1dB TP；
- c) 响度动态范围（LRA）不做硬性规定，但应避免因响度波动过大导致听众频繁调节音量。

附 录 A 空间音频有声剧制作流程
(资料性)

A.1 制作流程参考

为保障车载空间音频有声剧制作全流程的规范性与效率，建议遵循“前期筹备→核心制作→测试优化→交付归档”的阶段化流程，各环节需严格结合本标准中环境要求、声像定位、安全审查等核心条款开展，确保最终内容既满足空间音频的沉浸感，又符合车载场景的安全约束。

表1 制作流程总览表

阶段序号	阶段名称	核心任务节点	关联标准条款
1	前期筹备阶段	1.1 高保真素材筛选与预处理	5.1
		1.2 混音棚与监听系统校准	5.2
2	核心制作阶段	2.1 音乐/环境音效声床搭建	6.1
		2.2 声音对象（旁白 / 对白 / 动态音效）定位	6.2
		2.3 低频通道（LFE）控制	6.3
		2.4 响度校准	6.4
3	测试优化阶段	3.1 M1/M2多车型声场均匀度测试	6.5
		3.2 车载系统兼容性与播放稳定性测试	
4	交付归档阶段	4.1 ADM 母版与 AV3A 分发文件整理归档	7.1
		4.2 制作测试报告编制与提交	7.2
		4.3 原始素材与工程文件归档留存	

附录 B 人声/音效声场均匀度优化测试方法

(资料性)

本附录规定了车辆回放车载空间音频有声剧时，人声声场均匀度及音效声场均匀度的测试目的、准备工作、测量方法及数据记录。

B.1 测试目的

在安静状态下，测量车辆驾驶位、非驾驶位（副驾位、后排中间位）听到的有声剧人声（旁白+对白）、音效声压级；以此作为基准调整制作中各声道音量平衡，确保非驾驶位人声、音效声压级分别与驾驶位对应声压级差值在可控范围内，实现全座位人声、音效听感一致。

B.2 测试准备

- a) 车辆技术要求：应符合 T/UWA 009.3-4-2024 中第 6 节技术要求。
- b) 一般测试条件：应符合 T/UWA 009.3-4-2024 中第 7 节一般测试条件。

B.3 测试信号

- a) 综合音频信号：包含旁白+对白、音效，音乐的有声剧组合片段（持续片段 ≥ 30 秒），信号清晰无失真，应使用被测空间声官方渲染工具渲染为 7.1.4 声道 WAVE 文件，编码 LPCM，综合响度为 $-24\text{LKFS} \pm 2\text{LU}$ ， $\text{TP} < -1\text{dBTP}$ ；
- b) 单独人声信号：从综合片段中提取的典型人声信号，包含旁白和对白（人声有效持续片段应 ≥ 10 秒，宜 ≥ 30 秒），信号清晰无失真，应使用被测空间声官方渲染工具渲染为 7.1.4 声道 WAVE 文件，编码 LPCM，响度和真峰值与其在综合音频信号中保持一致；
- c) 单独音效信号：从综合片段中提取的典型音效信号（单一声效持续片段应 ≥ 3 秒，宜 ≥ 10 秒），信号清晰无失真，应使用被测空间声官方渲染工具渲染为 7.1.4 声道 WAVE 文件，编码 LPCM，响度和真峰值与其在综合音频信号中保持一致。

B.4 测量方法

- a) 音量校准：按照 T/UWA 009.3-4-2024 中 7.3.1 音频系统设置相关要求，播放综合音频信号，将驾驶位音量校准并锁定至适宜聆听声压级（ $\text{Leq}=75\text{dB(A)}$ ），确保播放流畅无中断
- c) 单独人声信号播放与测量：持续播放单独人声信号，依次在驾驶位、副驾位、后排中间位测点，分别测量单独人声信号的等效连续声压级，每个测点、每个项目重复测量 3 次，每次测量时长与被测文件长度保持一致。
- d) 单独音效信号播放与测量：持续播放单独音效信号，依次在驾驶位、副驾位、后排中间位测点，分别测量单独音效信号的等效连续声压级，每个测点、每个项目重复测量 3 次，每次测量时长与被测文件长度保持一致。
- e) 数据计算：分别计算各测点单独人声、单独音效信号的 3 次测量平均值，作为对应座位的实际声压级值
- f) 声压级均匀度测试方法与推荐参考：
 - 以驾驶位对应声压级为基准，对比副驾、后排中间位单独人声、单独音效信号声压级差值，参考以下范围开展声道参数优化，保障各位置听觉一致性：
 - i. 参考范围一：非驾驶位与驾驶位的声压级差值均 $\leq 3\text{dB}$ ，此状态下各位置听觉体验一致性极优；
 - ii. 参考范围二：非驾驶位与驾驶位的声压级差值 $\leq 6\text{dB}$ ，此状态下可满足常规听觉场景需求；
 - iii. 参考范围三：任意非驾驶位与驾驶位的声压级差值 $> 6\text{dB}$ ，建议重新调整空间音频有声剧的声道分配方案，以实现各位置音量平衡。

T/UWA XXXX—XXXX

B.5 数据处理要求

- a) 音量数据以 A 计权 L_{eq} 结果为准，保留 1 位小数
- b) 单次测量误差 $\leq \pm 1\text{dB}$ ，超出范围需重新测量