



中华人民共和国国家标准

GB/T XXXXX—XXXX

无障碍音视频出版物通用技术规范

General technical specifications for accessible audiovisual publications

（征求意见稿）

（2025-09-12）

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 发布

XXXX – XX – XX 实施

国家市场监督管理总局
国家标准化管理委员会 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 1

4 符号 2

5 总体要求 2

6 技术要求 2

 6.1 脚本撰写 2

 6.2 解说录制 3

 6.3 字幕制作 4

 6.4 声画合成 5

7 检测方法 6

 7.1 脚本质量检测 6

 7.2 解说录制质量检测 6

 7.3 字幕制作质量检测 6

 7.4 声画制作质量检测 7

附录 A（资料性） 无障碍音视频出版物视障人士脚本示例 8

附录 B（资料性） 无障碍音视频出版物听障人士脚本示例 9

参考文献 10

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。

本文件由全国新闻出版标准化技术委员会(SAC / TC 527)归口。

本文件起草单位：中国盲文出版社、中国传媒大学、山东电子音像出版社有限公司、北京臻鼎科技有限公司、上海音像教育出版社有限责任公司、中国电影科学技术研究所（中宣部电影技术质量检测所）、人民教育电子音像出版社有限公司、北京时代华文书局有限公司、中国聋人协会、中国新闻出版研究院、北京甲骨今声科技有限公司、百视通网络电视技术发展有限公司、湖南电子音像出版社有限责任公司、中国科学技术大学出版社有限责任公司。

本文件主要起草人：

引 言

为进一步加强中国无障碍音视频出版物制作技术在我国公共文化服务领域的普及与推广，本文件在T/CSMPTE 13—2021《无障碍影视节目制作技术规范》的基础上，结合团体标准实施情况和视听障碍人士对无障碍音视频出版物制作技术的新需求，尤其对听障人士字幕制作技术进行更科学更全面地阐述和表述，从而规范中国无障碍音视频出版物在脚本撰写、解说录制、字幕制作、声画合成等方面的制作技术内容，提高中国无障碍音视频出版物技术人员的服务和制作水平，最大程度地满足中国视听障碍人士对无障碍音视频出版物的精神文化需求，制定和实施更具权威性、覆盖范围更广的推荐性的《无障碍音视频出版物通用技术规范》国家级标准。

无障碍音视频出版物通用技术规范

1 范围

本文件规定了为视、听障碍两类人群制作无障碍音视频出版物的总体要求、技术要求，描述了检测方法。

本文件适用于视、听障碍人士欣赏的无障碍音视频出版物的制作，其他无障碍音视频作品的制作参照使用。

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。

GB/T 16463—1996 广播节目声音质量主观评价方法和技术指标

GB/T 33665—2023 数字出版物声频视频技术要求及检测方法

GY/T 168-2001 广播音频数据文件格式规范 广播波形格式（BWF）

3 术语和定义

下列术语和定义适用于本文件。

3.1

无障碍音视频出版物 accessible audiovisual publication

在原音视频出版物基础上增加视障人士解说、说明性字幕，可供视听障碍人士欣赏的音视频出版物。

3.2

无障碍音视频出版物脚本 script for accessible audiovisual publication

根据视听障碍人士欣赏音视频出版物的需求，在原音视频出版物基础上撰写剧情画面解说、说明性字幕的文稿。

3.3

无障碍音视频出版物视障人士脚本 script for visually impaired users in accessible audiovisual publication

根据视障人士欣赏无障碍音视频出版物需求，在原音视频出版物基础上，将必要的画面内容转化为适合视障人士的文字解说的脚本。

3.4

无障碍音视频出版物听障人士脚本 script for deaf and hard-of-hearing People in accessible audiovisual publication

根据听障人士欣赏无障碍音视频出版物需求，在原音视频出版物基础上，将必要的声音内容转化为适合听障人士的字幕的脚本。

3.5

视障人士解说 audio description

为视障人士提供的，对没有对白的音视频出版物画面进行的解说。

3.6

说明性字幕 caption

为听障人士或有特殊需要的观众理解音视频出版物中的对白、背景声音、非声音内容加以解释的文字。

注：非声音内容包括但不限于外文的中文翻译（语音翻译、画面中外语元素的非声音内容翻译）等。

4 符号

下列符号适用于本文件。

♪：用于音乐的提示符号。

♪：用于歌词字幕的提示符号。

【】：用于描述除音乐及歌词外背景声音的提示符号。

5 总体要求

无障碍音视频出版物应按照脚本撰写、解说录制、字幕制作、声画合成四大环节，规范有序地开展制作工作，确保无障碍音视频出版物脚本撰写细致、录音播讲严谨、字幕匹配得当、技术合成专业，且符合视听残障人士欣赏音视频出版物的需求。

6 技术要求

6.1 脚本撰写

6.1.1 脚本撰写内容

脚本撰写内容应符合以下技术要求：

——无障碍音视频出版物视障人士脚本应当对片名、国别、背景、故事梗概、主要演职人员等信息进行简要介绍。音视频作品结束时应进行相应的结束提示，应对音视频作品版权授予方给予鸣谢。若音视频作品涉及外文，应翻译成中文。对没有声音提示的时间、场景、人物、主要事件以及其他附加信息进行描述。脚本撰写应充分遵循客观性、通用性和准确性等原则，用词应简练、准确，不得干扰音视频出版物对白，避免使用口语、专业术语和俚语方言。对于无法避免的关键专业术语，应在首次出现时进行简洁明了的解释。

——无障碍音视频出版物听障人士脚本应按照时间点插入对音效和音乐的提示性描述，描述应简洁、准确。对非声音内容的说明性字幕应做到准确。

6.1.2 脚本撰写格式

脚本撰写格式应符合以下技术要求：

- 无障碍音视频出版物视障人士脚本撰写时，应以“各位观众，现在播放的是……”开头，对音视频出版物内容、影片主创进行介绍。对音视频出版物内容进行解说时，应先标明起始时间码或标明台词起始关键词。影片剧情播放结束时，应对影片版权授予方给予鸣谢。无障碍音视频出版物视障人士脚本具体示例见附录 A。
- 无障碍音视频出版物听障人士脚本撰写时，应标明音乐起止时间码或音效起止时间码，以及无障碍音视频出版物听障人士脚本具体内容。无障碍音视频出版物听障人士脚本具体示例见附录 B。

6.2 解说录制

6.2.1 解说

解说应符合以下技术要求：

- 无障碍音视频出版物解说人员的普通话水平应不低于一级乙等或同等水平，应按照脚本进行解说，不应干扰音视频出版物正常对白，解说语气语调应与音视频出版物风格相一致。
- 若视频涉及外文需要翻译成中文，应采用区别于无障碍音视频出版物视障人士脚本的另一种声音形式进行解说。

6.2.2 录音棚技术

录音棚技术指标应符合表1和表2的规定。

表 1 录音棚建筑声学技术要求

序号	项目	技术要求	
1	对白录音棚、录音控制室背景噪声	对白录音棚	所有必要工作设备开启时，动态背景噪声应不高于 NR25 噪声评价曲线。
		录音控制室	所有必要工作设备开启时，动态背景噪声应不高于 NR25 噪声评价曲线。
2	对白录音棚、录音控制室空气声计权隔声量（ $D_{nT,w}$ ）	$\geq 65\text{dB}$	
3	对白录音棚混响时间（RT）	$0.032808V^{0.333333} \leq RT \leq 0.07653V^{0.287353}$ （测量频率为 500Hz，单位为秒（s），V 为观众厅实际容积，单位为立方米（ m^3 ），也可按容积从图 1 中查得 RT 上、下限。）	
4	录音控制室混响时间（RT）	$0.032808V^{0.333333} \leq RT \leq 0.07653V^{0.287353}$ （测量频率为 500Hz，单位为秒（s），V 为观众厅实际容积，单位为立方米（ m^3 ），也可按容积从图 1 中查得 RT 上、下限。）	
5	对白录音棚混响时间频率特性	混响时间频率特性曲线应保持平滑，倍频段频率特性的允差范围应符合表 2 中对白录音棚的规定。	
6	录音控制室混响时间频率特性	混响时间频率特性曲线应保持平滑，倍频段频率特性的允差范围应符合表 2 中对录音控制室的规定。	

序号	项目	技术要求
7	走廊和对白录音棚之间门空气声隔声量	$\geq 55\text{dB}$
8	走廊和录音控制室之间门空气声隔声量	$\geq 35\text{dB}$

表 2 电影对白录音棚、录音控制室混响时间倍频程段频率特性要求

类别	中心频率							
	63Hz	125Hz	250Hz	500Hz	1000Hz	2000Hz	4000Hz	8000Hz
对白录音棚	宜 1.0~1.4	1.0~1.2	1.0~1.2	1.0	0.8~1.0	0.8~1.0	0.7~1.0	宜 0.6~0.9
录音控制室	宜 1.0~1.4	1.0~1.2	1.0~1.2	1.0	0.8~1.0	0.8~1.0	0.7~1.0	宜 0.6~0.9

6.2.3 解说录制指标

解说录制指标应符合表3的规定。

表 3 解说录制指标

类别	指标要求
音质	声音真实、层次分明，无异常杂音
内容	与脚本内容相符，不应出现错读、漏读，零口误
语速	语速以每分钟180~240字为宜，上下文段落停顿等待时间以0.5 s~1 s为宜
情感表达	根据情景的刻画，解说语气语调应与音视频出版物风格相一致
语义表述	根据情景的刻画，语义表述准确、客观，音节清晰
音量控制	无削波失真，且不应过大或过小，人声与背景音乐比例调节得当，不同录音者或不同影片音量保持均衡
话筒角度	解说人员应正对录音话筒，主轴偏离角度宜不大于 $\pm 30^\circ$
录制音频音量	录音音量宜在[-18 dBFS,-6 dBFS]区间内，峰值电平不大于-2 dB

6.2.4 音频文件格式

音频文件格式要求应符合表4的规定。

表 4 音频文件格式要求

类别	指标要求	
文件格式	应符合GY/T 168—2001的规定	
文件类型	无压缩WAV，FLAC，APE，ALAC，WavPack(WV)格式	
采样频率	44.1 kHz或48 kHz	
量化深度	16 bit或24 bit	
文件命名方式及声道映射	2 声道立体声	Track_L: 左声道
		Track_R: 右声道
	5.1.2 环绕立体声	Track_L: 左声道

类别	指标要求	
		Track_R: 右声道
		Track_C: 中声道
		Track_SW: 次低音声道
		Track_Ls: 左环绕声道
		Track_Ls: 左环绕声道
		2个天空声道（如前置头顶或反射式扬声器）

6.3 字幕制作

6.3.1 字幕要求

字幕应符合以下技术要求：

- 说明性字幕应与相应的音视频出版物声源、非声音内容等的时间点一致，不应提前或延迟。
- 字幕应准确无误，无错别字、无漏字、无多字现象且不应影响对白字幕。
- 字幕中不允许加任何与无障碍音视频出版物无关的标识、角标及移动字幕。
- 说明性字幕中描述音乐的字幕需在开始处应用两个八分音符“♪”进行标识，用于描述背景声音的提示符号应用方括号“【】”进行标识；用于歌词字幕的提示符号应用一个八分音符“♪”进行标识。
- 多人对话时，宜在对白字幕前以冒号加以角色标注，且对白字幕不宜超过两行，参见附录B。

6.3.2 对齐方式

字幕行高应紧贴字幕安全框内框，不压原片音视频出版物字幕。如果字幕是对背景声的说明，应水平居中对齐；如果字幕是对画面特定元素的说明，字幕内容应位于需说明元素附近便于理解且最大限度不遮挡画面其他元素的位置，不要求水平居中对齐。

6.3.3 字体

字幕宜为黑体。

6.3.4 颜色

颜色宜符合以下技术要求：

- 字体颜色宜为白色，RGB值为（255、255、255）
- 字体描边宜为黑色，RGB值为（0、0、0），宽度4，透明度100（完全不透明）。

6.3.5 位置

歌词宜置于画面上方且居中。描述背景声音的字幕宜在对白字幕上方单独行居中。若对白是外语，则应在对白字幕上方单独行居中加外国语言类型解释，参见附录B。

6.3.6 分行处理

字幕应分行断句合理，气口恰当，合乎剧情；字幕时间间隔应参考描述内容需要，一般两句字幕间应至少有2帧间隔。

6.3.7 字幕字数

每屏字幕字数不宜超过14个字（含特殊符号）。

6.3.8 字幕时长

每屏字幕显示时长宜3 s~5 s。

6.3.9 字幕出现方式

字幕的字号以及大小等呈现形式宜与原音视频出版物的风格和字幕的形式相匹配。

6.4 声画合成

6.4.1 音频编辑技术

音频编辑应符合以下技术要求：

- 无障碍音视频出版物以原音视频出版物声音为主，解说为辅，同时解说不应超过音视频出版物原声的音量。通常情况下宜小于原音视频出版物音量。当音视频出版物原声音量过大或过小时，应适当调整原声音量，使之与配音音量相匹配，确保清晰听见配音内容。背景噪声电平应小于-50 dBFS。
- 声音层次分明，音量保持均衡，无异常杂音，无削波失真。

6.4.2 视频编辑技术

视频编辑应符合以下技术要求：

- 新建的视频工程文件/事件（根据不同编辑软件名称不同）参数应保持与原片一致或不低于原片水平。
- 工程文件/事件分辨率宜设置为1920×1080，即宽高比为16:9。
- 分辨率为960×720、720×576或宽高比为4:3的音视频出版物，导入工程文件/事件中时，应当对视频采取增加遮幅等方式，补齐至16:9的视觉效果。
- 帧速率根据原片情况可设定为24 fps、25 fps、30 fps或50 fps。音频通道为立体声通道，采样速率为48 kHz。
- 分辨率为2048×1080或2048×858的音视频出版物，导入工程文件/事件中时，宜保留原宽高比。
- 色度采样宜不低于YUV 4:2:2。
- 色深宜不低于8 bit。
- 图像应色彩自然统一、稳定、画质清晰、无变形、无杂乱信号（如闪烁、偏色、波纹、花屏、噪波等现象），图像中不可出现彩条、夹帧、白屏等现象。

6.4.3 成片输出技术

成品输出应符合以下技术要求：

- 音视频出版物原片视频格式应为MP4，分辨率为1920×1080，宽高比为16:9，帧速率根据原片情况可设定为24 fps、25 fps、30 fps或50 fps，编码形式为H.264。
- 解说音频文件格式宜为无压缩WAV、FLAC、APE、ALAC、WavPack(WV)格式。
- 字幕文件格式宜为.srt或.ass。
- 色度采样宜不低于YUV 4:2:2。
- 色深宜不低于8 bit。
- 视频码率宜不低于23 Mbps。

7 检测方法

7.1 脚本质量检测

无障碍音视频出版物视障人士脚本及无障碍音视频出版物听障人士脚本应按6.1的要求进行检测。

7.2 解说录制质量检测

解说录制质量应按GB/T 33665—2023中6和GB/T 16463—1996中5~7的要求进行检测。

7.3 字幕制作质量检测

字幕制作应按6.3的要求进行检查。

7.4 声画制作质量检测

声画制作质量应按6.4的要求进行检测。

附录 A

(资料性)

无障碍音视频出版物视障人士脚本示例

大明劫

各位观众，现在播放的无障碍音视频作品是《大明劫》。该影片讲述了崇祯十五年（1642年）瘟疫横行，李自成攻打开封，明朝江山内忧外患，孙传庭与吴又可分别誓死抗击闯军与瘟疫的故事。最后瘟疫终于被控制住，但是国家的病症能否被战争治好吴又可和孙传庭都无法预料。该片是由王竞指导，谢晓东制片，影片主要人物吴又可由冯远征饰演，孙传庭由戴立忍饰演，冯氏由冯波饰演，赵川由钱学格饰演，余乔迁由司源饰演，顾清远由马精武饰演，任琦由胡晓光饰演，崇祯皇帝由余少群饰演。

00:01:33 画面出现字幕：1642年，距明朝覆灭还有两年，李自成率领农民军包围开封。明朝大厦将倾。

00:01:44 画面出现片名：大明劫

00:01:53 开封城楼上插着明字旗，乱箭从城楼射出。闯军如水银泻地般围拢，架起云梯准备攻城。

00:02:12 随着号令兵的喊声和旗语，士兵们将城楼上的炮口抬起，向城外发射。

00:02:19 闯军发起猛烈地攻击。

00:02:33 城内凌乱不堪。

00:02:39 城垛口一名持箭向外瞄准的士兵中箭后倒地。

00:02:52 士兵们奋力防守。

00:02:58 炮弹在闯军中炸响。

00:03:08 一个穿着甲胄的明军将领，正从箭垛用单筒望远镜向外看。

00:03:20 一小队闯军停止冲锋，有序撤退。

从望远镜中看去，远处，李自成的军队扛着一面面闯字旗，正在从四面八方聚集，要发动一轮更大更强的攻击。

00:03:37 士兵将一张写有“开封告急”的纸条折好，放到信鸽脚环儿里，放飞信鸽。

00:03:48 北京刑部大牢内，太监总管王承恩走来。狱卒打开一间牢房的牢门。

00:04:28 墙上的潼关地图隐约可见。

00:04:30 紫禁宫文华殿

.....

01:45:28 孙传庭战死于潼关，冯氏投井自杀，孙传庭与冯氏合葬于山西下花庄村，其陵墓于上世纪60年代文革浩劫中被损毁。

01:45:39 吴又可领先世界近200年开创性地提出了病毒学说及病毒的传播方式。崇祯15年（公元1642年），吴又可回到老家苏州东山，于1643年完成了不朽名著《瘟疫论》，从而奠定了中医治疗传染病的理论基础，书中所记经方“达原饮”用于治疗非典收到了奇效。

《大明劫》放映完毕，谢谢收看。衷心感谢电影频道节目制作中心、江苏众道影业投资有限公司、中央新闻纪录电影制片厂（集团）对中国无障碍音像出版事业的大力支持。

附录 B

(资料性)

无障碍音视频出版物听障人士脚本示例

龙之战

00:00:40 【打雷声】
 00:00:59-00:01:35 ♪ 紧张
 00:01:10 【马蹄声】
 00:01:46 【马鸣声 水下气泡声】
 00:02:03 【鹰叫声】
 00:02:08-00:03:45 ♪ 紧张 恐怖
 00:02:41 【枪声】
 00:03:13 【马惊叫声】
 00:03:23 【人呼喊声】
 00:03:46 【鹰叫声】
 00:03:53 【马蹄声】
 00:04:02 【法语】
 00:04:08 【法语】
 00:04:22 【开锁声】
 00:04:42 【鞭子抽打声】
 00:05:12 【士兵呼嚎声】
 00:05:31 【鹰叫声】
 00:05:26-00:06:46 ♪ 紧张
 00:05:32 【法语】
 00:05:45 【法语】
 00:05:47 【法语】
 00:06:21 【法语】
 00:06:23 【法语】
 00:06:51 锁链响声
 00:07:02 【法语】
 00:07:04 【枪响声】
 00:07:10 【炮火声】
 00:07:12-00:07:57 ♪ 缓慢 低沉

 01:42:56-01:44:03 ♪ 抒情
 01:44:30-01:47:12 ♪ 浑厚 宏伟
 01:44:30 ♪ 落日在风中飘摇
 01:44:35 ♪ 我站在这块土地上
 01:44:43 ♪ 天涯海角连着四面八方
 01:44:47 ♪ 丹心照汗青不只一个人在想
 01:44:49 ♪ 我的国土我的声音回荡

参 考 文 献

- [1] GB/T 7400—2011 广播电视术语
 - [2] GB/T 7401—1987 彩色电视图像质量主观评价方法
 - [3] GB/T 16463—1996 广播节目声音质量主观评价方法和技术指标要求
 - [4] GB/T 31001—2014 高清晰度数字电视主观评价用测试图像
 - [5] GB/T 31219.5—2016 图书馆馆藏资源数字化加工规范 第5部分：视频资源
 - [6] GB/T 33665—2023 数字出版物声频视频技术要求及检测方法
 - [7] GB 50763 无障碍设计规范
 - [8] GY/T 168-2001 广播音频数据文件格式规范 广播波形格式（BWF）
 - [9] GY/T 288—2014 影视节目送审数字母版字幕技术要求
 - [10] GY/T 301—2016 视频节目对白字幕数据格式规范
 - [11] YD/T 3543—2019 视障者多媒体信息处理技术要求
 - [12] T/CSMPTE 13—2021 无障碍影视节目制作技术规范
 - [13] T/CSMPTE 9—2020 电影鉴定放映室和录音棚技术要求和测量方法
-