

国家公共文化数字支撑平台 数字资源标准规范

第三部分 数字资源加工格式规范

委托方：文化部全国公共文化发展中心

研制方：北京大学

2014 年 6 月

目 录

1	编制说明.....	1
2	范围.....	1
3	规范性引用文件.....	1
4	术语和定义.....	2
4.1	模拟视频 (Analog Video)	2
4.2	数字视频 (Digital Video)	2
4.3	采集 (Capturing)	2
4.4	视频分辨率 (Video Resolution)	2
4.5	比特率 (Bit Rate)	2
4.6	视频帧速 (Frame Rate)	2
4.7	音频采样 (Audio Sample)	2
4.8	压缩 (Compression)	2
4.9	编码 (Coding)	3
4.10	盘 (Disc)	3
4.11	磁带 (Magnetic Tape)	3
4.12	光盘 (Optical Disc)	3
4.13	存储 (Storage)	3
4.14	存储卡 (Memory Card)	3
4.15	发布服务级 (Service Level)	3
4.16	编辑保存级 (Preservation Level)	3
4.17	格式 (Format)	4

4.18 文件扩展名 (File Extension)	4
4.19 F4V.....	4
4.20 高级音频编码 (Advanced Audio Coding; AAC)	4
4.21 MP4.....	4
4.22 素材交换格式 (Material Exchange Format; MXF)	4
4.23 线性脉冲编码调制 (Linear Pulse Code Modulation)	4
4.24 JPEG 2000.....	4
4.25 压缩比.....	5
5 资源级别、加工标准和格式体系	5
5.1 编辑保存级.....	5
5.2 发布服务级.....	6
6 数字加工流程.....	7
6.1 加工准备.....	7
6.2 内容采集.....	8
6.3 压缩转换.....	9
6.4 视频编辑.....	9
6.5 元数据加工.....	9
6.6 质量检验.....	11
6.7 过程管理.....	11
7 命名规则.....	11
参考文献.....	13

1 编制说明

《国家公共文化数字支撑平台视频资源加工规范》是“国家公共文化数字支撑平台资源建设标准规范研制项目”的研究成果之一，是为创建国家公共文化数字支撑平台服务的。本规范是建立在调研国内外数字资源加工规范的发展现状及相关标准、分析数字资源和应用环境类型的基础上，为在新兴网络技术和新媒体技术快速发展的背景下保障国家公共文化数字支撑平台数字资源的有效整合、管理和发布制定的。

2 范围

《国家公共文化数字支撑平台视频资源加工规范》规定了视频资源数字化加工应遵循的标准，包括工作流程标准、元数据加工标准、视频数字对象的命名规则等。

本规范给出了适用于视频资源发布服务级的推荐标准。

本规范用于模拟视频、数字视频的采集、编码转换、编辑以及发布。

本规范适用于全国各级、各类型文化机构在模拟视频、数字视频的采集、编码转换、编辑以及发布工作中使用，同时也可供从事相关业务的单位参考使用。

3 规范性引用文件

下列文件对于本规范的应用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本规范。凡是不注日期的引用文件，其最新版本（包括所有的修改版）适用于本规范。

GB/T 7400-2011 广播电视术语

GB/T 14857-1993 演播室数字电视编码参数规范

GY/T 155-2000 高清晰度电视节目制作及交换用视频参数值

ITU-T 建议 H.264(2013) 通用视听业务的先进视频编码 (Advanced video coding for generic audiovisual services)

ISO/IEC 15444-1:2004 标准 JPEG 2000 图像编码系统：核心编码系统 (JPEG

2000 image coding system: Core coding system)

4 术语和定义

下列术语和定义适用于本文件。

4.1 模拟视频 (Analog Video)

以电磁信号记录的视频资料。

4.2 数字视频 (Digital Video)

以数字信息记录的视频资料。

4.3 采集 (Capturing)

将模拟的视频内容采样编码为数字内容。

4.4 视频分辨率 (Video Resolution)

视频图像的大小或尺寸。

示例：355 × 288, 176 × 144, 640 × 480, 1024 × 768。

4.5 比特率 (Bit Rate)

传送二进制的速度。单位为每秒传送的位数。

4.6 视频帧速 (Frame Rate)

单位时间内传输的视频画面数量。一般用帧/秒或 fps 表示。

4.7 音频采样 (Audio Sample)

视频数据中音频信号的采样方式。

示例：立体声 48kHz

4.8 压缩 (Compression)

以较少的位表达原始数据的编码过程。

4.9 编码 (Coding)

数据变换或数据表示的过程。

[GB/T 17975.2:2000 术语和定义 3.20]

4.10 盘 (Disc)

信息被压制或压缩在材料表面上，可以用特殊的设备将其恢复的，扁平的圆形数据媒介。

4.11 磁带 (Magnetic Tape)

通常用塑料制成、表面覆盖着可以记录数据的磁化层的带子。

4.12 光盘 (Optical Disc)

通过激光扫描仪记录、回放或检索数据的盘（4.10）。

4.13 存储 (Storage)

为数据和文献的后续利用而进行的放置和保持的过程。

4.14 存储卡 (Memory Card)

一种固态电子快闪存储器数据存储设备，多为卡片或者方块状。

4.15 发布服务级 (Service Level)

以获取和使用为目的，发布在服务器上，用户能够通过因特网快速下载、浏览的数字视频。这种级别的文件通常分辨率较低、体积小。

4.16 编辑保存级 (Preservation Level)

以编辑保存为目的，在确保数字视频数据的逻辑和内容完整性的前提下，编辑并保存视频数据，使得随着时间的流逝和技术的变化，还能够对存档的数据进行读取和编辑。

4.17 格式 (Format)

预先确定的数据媒介上的数据排列。

[GB/T 4894-2009 术语 4.1.1.4.9]

4.18 文件扩展名 (File Extension)

计算机文件的末尾显示的字母，帮助计算机区分文件的格式。

示例：.f4v，.mp4，.avi。

4.19 F4V

Flash 视频 (Flash Video; FLV) 格式的扩展，支持 H.264 编码。

4.20 高级音频编码 (Advanced Audio Coding; AAC)

由 Fraunhofer IIS、杜比实验室、AT&T、Sony 等公司共同开发的一种音频压缩标准，广泛用于各种音频设备。

4.21 MP4

一种压缩标准，使用 MPEG-2 AAC 技术，以储存数码音频及数码视频为主。

4.22 素材交换格式 (Material Exchange Format; MXF)

美国电影电视工程师协会 (Society of Motion Picture and Television Engineers; SMPTE) 组织定义的一种专业音视频媒体文件格式。MXF 文件通常被视为一种“容器”文件格式。

4.23 线性脉冲编码调制 (Linear Pulse Code Modulation; LPCM)

以连续线性取样的方式撷取模拟信号转换成数字信号。

4.24 JPEG 2000

由联合图像专家组 (Joint Photographic Experts Group; JPEG) 创建和维护的一种基于小波变换的图像压缩标准，同时支持有损数据压缩和无损数据压缩。

4.25 压缩比

数据被压缩的比例，输入流比特率与输出流比特率的比值。

5 资源级别、加工标准和格式体系

5.1 编辑保存级

编辑保存级的数字视频资源主要用于进行编辑和长期保存，也可作格式转换和复制的母本，以备必要时再利用，要求有互操作性强、使用广泛、质量高的特点，因此推荐采用 MXF 封装的 JPEG2000 无损压缩编码，也可在存储和应用成本有限的情况下使用 MP4 封装的 H.264 编码作为二级替代标准。

根据不同存储和应用成本推荐的规格参数，如表 1.1 和表 1.2。

表1.1 编辑保存级级视频资料一级加工标准和格式体系

存储 与应 用级 别	主要编码参数				
	视频分 辨率	视频帧速 (fps)	视频编码	封装格式	音频编码
一级	与视频 源保持 一致	与视频源保持一 致	JPEG 2000 (整数 5/3 小波 变换)	MXF (OP1a)	LPCM

国家公共文化数字支撑平台的视频来源和介质多种多样，视频源参数也不尽相同，本规范中以分辨率为标准PAL (720 × 576) 和HDTV (1920 × 1080) 的视频源为例示明编辑保存级级视频资料二级加工标准和格式体系。

表1.2 编辑保存级级视频资料二级加工标准和格式体系

存储 与应 用级 别	主要编码参数						
	视频分 辨率	视频 帧速 (fps)	视频比 特率 (kbps)	音频比 特率 (kbps)	视频编码	封装 格式	音频编码
二级	720 × 576 (PAL)	25/30	3000	384	H.264 (Hi422P@Level3.1 gop=25, ref=8, crf=18)	MP4	AAC, 立体声, 48kHz, 16bit 量化

					8 比特量化)		
	1920 × 1080 (HDTV)		15000		H. 264 (Hi422P@Level4.2 gop=25, ref=8, crf=18 8 比特量化)		
注：若视频源参数不同于表中所列，则应保持其各项参数不变，以避免上下变换、重采样等影响编辑保存级资源的质量，从而做到最大程度的无损转换与编码压缩，同时视音频比特率也应相应变化，在存储与应用成本允许的情况下就高不就低；推荐使用 x264 编码器进行 H. 264 的视频编码							

5.2 发布服务级

根据国家公共文化数字支撑平台视频源质量，将原始资源分为两类，第一类为高清视频资源，这类视频资源的画面宽高比一般为 16:9，多来自于广播级高清摄像机拍摄形成，介质多为存储卡、数字视频光盘和高清磁带；第二类为非高清视频资源，这类视频资源的画面宽高比一般为 4:3，来源和介质多种多样。

发布服务级的数字视频资源主要用于检索、浏览、下载等服务，根据具体用途和使用对象的不同又可分为以下 3 个级别：

（1）高级：多用于投影仪、计算机、移动播放器、共享机、非编设备以及电子政务外网（或硬盘方式）下发

（2）中级：多用于局域网、卫星广播及投包、卫星播发以及政务外网（或硬盘方式）下发。

（3）低级：多用于互联网访问。

综上，根据不同视频源质量以及使用目的和应用场景推荐的规格参数，如表 2。

表2 发布服务级视频资料加工标准和格式体系

画面 宽高 比	应用 级别	主要参数建议						
		视频分辨率	视频帧 速 (fps)	视频比 特率 (kbps)	音频比 特率 (kbps)	封装 格式	视频编码	音频编码

16: 9	高级	1280 × 720 (720P)	25/30	3200	256	MP4 或 F4V	H. 264 (HiP@Level3. 1, gop=25, ref=5, 8 比特量化, 4: 2: 0 取样结构)	AAC, 48kHz, 立体声, 16bit 量化		
	中级	854 × 480 (FWVGA)		1400	128		H. 264 (HiP@Level3, gop=25, ref=3, 8 比特量化, 4: 2: 0 取样结构)	AAC, 44. 1kHz, 立体声, 16bit 量化		
	低级	640 × 360 (nHD)		800						
4: 3	高级	720 × 576 (PAL)		1500	128		H. 264 (HiP@Level3. 1, gop=25, ref=5, 8 比特量化, 4: 2: 0 取样结构)	AAC, 48kHz, 立体声, 16bit 量化		
	中级	640 × 480 (VGA)		1000					H. 264 (HiP@Level3, gop=25, ref=3, 8 比特量化, 4: 2: 0 取样结构)	AAC, 44. 1kHz, 立体声, 16bit 量化
	低级									

注：帧速以视频源帧速为上限，并且以忠于视频源帧速为最优；若资料来源音频为单声道，则不必加工为立体声；若非高清视频源的分辨率低于 VGA 级别，则以其原分辨率作为各应用级别的分辨率；建议根据视频源内容决定使用可变码率或固定码率；本表所推荐的加工标准与格式体系同样适用于移动设备访问；推荐使用 x264 编码器进行 H. 264 的视频编码

6 数字加工流程

6.1 加工准备

- 应对拟加工的视频资源保存情况进行评估、盘点、核查，筛选出需要加工的视频资源并进行登记，对有缺损和不适合加工的资源登记后进行修复或保护等处理，必要时对视频资源的载体进行适度的清洗和修复。
- 应在加工之前对加工级别及技术参数、采集与处理、元数据加工、命名、

质量管理等进行一致性约定。

6.2 内容采集

6.2.1 模拟视频采集

6.2.1.1 采集方式

通过视频采集卡将模拟磁带内容按采集标准采集到存储设备。

6.2.1.2 采集内容

磁带首尾导带之间的所有内容。

采集中因不可抗拒的原因被迫中断且无法修复的，保存已采集视频，并将内容信息记录到元数据。

6.2.2 数字视频采集

6.2.2.1 采集方式

(1) 数字磁带

通过视频采集卡和数据线将数字内容按采集标准采集到存储设备。

(2) 数字视频光盘或存储卡

直接通过软件或读卡设备将视频内容及关联文件拷贝到计算机，或通过格式转换软件将其转换成选定的采集格式。

6.2.2.2 采集内容

(1) 数字磁带

磁带首尾导带之间的所有内容。

采集中因不可抗拒的原因被迫中断且无法修复的，保存已采集视频，并将内容信息记录到元数据。

(2) 数字视频光盘和存储卡

内容包括主视频文件、片花、花絮、字幕。

采集中因不可抗拒的原因被迫中断且无法修复的，保存已采集视频，并将内容信息记录到元数据。

含多语种字幕的，保留所有语种和字幕文件，并将其与视频主文件一起封装关联保存。

6.3 压缩转换

选择需加工的资源，通过相应设备转换成目标格式。

6.4 视频编辑

对拟提供服务的数字对象，如需进行处理，按下列情形进行：

——影像中的黑屏、镜头严重抖动、因操作失误造成的无关影像等在不影响节目内容判读的原则下进行删除处理，以符合浏览惯例。

——影像中加插的广告等不剪切。

——影像的色彩饱和度、亮度和对比度影响视频效果的，需进行色彩饱和度、亮度、对比度处理，以达到更为清晰的视频效果。

——不对数字成品作音频方面编辑加工。

——对自建资源添加片头和片尾文字信息。片头包含：活动主题、主讲人、主持人；片尾包含：活动主办/协办单位、收藏单位、拍摄时间。片头片尾字幕底色为蓝色或黑色，字为白色。

——字幕挂接。字幕的出现和消失需与原始视频中字幕的出现和消失在时间线上保持一致。字幕含多语种的，需挂接多语种字幕。

6.5 元数据加工

视频数据加工过程中涉及的元数据的创建应遵守《国家公共文化数字支撑平台数字资源元数据规范》项目成果的规定。著录人员在元数据加工环节中需要创建描述元数据，在视频数据加工过程中需要创建的管理元数据如下表 2。

表2 视频资源数字化加工过程中创建的管理元数据

元素	定义
标识符	唯一识别信息资源的标识
视频编码格式	视频数字化的编码方式
视频取样格式	视频信息编码时从原始图像信息中取样的方式
视频数据码率	视频数据码流每秒传输的比特数
音频编码格式	数字音频采样序列的编码格式
音频采样频率	录音设备在一秒钟内对声音信号的采样次数，如

元素	定义
	44.1KHz
音频数据码率	音频数据码流每秒传输的比特数
音频位深度	在视频资源数字化时声音样本信息使用的比特数，也称为声音样本量的量化深度，如 16 位
声道格式	描述视音频资源的声音通道之间的关系，如立体声
色彩	描述视频资源显现的色彩种类，如黑白
画面宽高比	视频资源显示画面的宽度与高度之比，如4：3
时长	播放资源正式有效内容的实际时间长度
分辨率	视频图像的大小或尺寸
字幕形式	视频资源的字幕表现形式
声音特征	视频资源的声音特征
文件格式	资源的数据格式，用于注明需要什么软件或硬件来显示和执行此文件
文件大小	视频资源文件的大小，如 150M
语种	资源的语种
来源	与当前资源来源有关的资源
关联	指明与本资源存在某种关系的其他资源
源载体	当本资源转换自其他载体时，说明资源的源载体特征
加工单位	资源的加工机构
加工类型	包括内容采集、压缩转换、视频编辑等
加工时间	加工操作发生的时间
加工设备	加工的硬件环境，如“大洋”非线性编辑系统
注：表中的元素及其定义来自并扩展于《国家公共文化数字支撑平台数字资源元数据规范》；表中的元素如果不能满足实际加工需要，可以根据《国家公共文化数字支撑平台数字资源元数据规范》的规定添加或自定义扩展元素项。	

6.6 质量检验

加工过程中应对视频数字对象、视频对象元数据进行质量检验，确保完全合格。

视频数字对象的检验项目，包括视频文件夹和文件名命名正确，文件采集完整、无漏帧，所有文件能正常打开、播放和显示，分辨率、采样频率和其他技术指标达到“5 资源级别、加工标准和格式体系”中的技术规范。

保证视频正确反映原始资源。对于出现问题的视频，确定其为采集过程中造成的，应进行重新采集。

6.7 过程管理

应建立过程管理表，详细登记视频数字加工流程，并及时整理、汇总，相关文档装订成册，在数字化工作完成的同时建立起完整、规范的记录。

应建立视频数字加工各环节的安全管理机制，确保实物在加工过程中不丢失、不损坏。

7 命名规则

应为加工后的视频数据文件或文件夹及存储介质进行命名。推荐的命名规则由以下部分组成：资源代码+资源级别代码+资源名。各部分直接连接，不使用连接符号。

各部分的说明如下：

（1）资源代码为资源编号+加工年份。资源编号代表资源的 ID，加工年份为 2 位数字。

（2）资源级别代码由两位字母组成，代表资源的应用级别，在本规范中用字母 A 表示编辑保存级，因存储和应用成本不同有分为一级和二级，分别用字母 H 和 L 表示；用字母 D 表示发布服务级，因用途和使用场景不同又分为高级、中级和低级。

高级：用字母 DH 表示。

中级：用字母 DM 表示。

低级：用字母 DL 表示。

(3) 资源名的命名规则应符合，但不限于以下约束：

- 拥有唯一标识符、具备连续一致的结构，不能与其它资源标识符重复；
- 应严格遵守计算机系统对文件命名的限制；
- 具备长期可用性。文件命名方式不依赖于某种处理或者系统。文件名包含的信息不应随着时间的推移而改变；
- 文件扩展名采用三位半角小写字母；
- 服务文件的命名应与主文件的命名保持一定的继承性。

参考文献

- [1] 朱强等. 国家图书馆视频数据加工标准和操作指南[M]. 北京: 国家图书馆出版社, 2011
- [2] 张成昱, 周虹, 曾婷, 等. 我国数字图书馆标准规范建设: 通用数字资源(视频数据)格式标准分析报告. [EB/OL]. [2014-03-08]. <http://cdls.nstl.gov.cn/cdls2/w3c/2003/Process/>
- [3] 孙一钢, 龙伟, 赵四友. 我国数字图书馆标准规范建设: 数字资源加工标准研究报告. [EB/OL]. [2014-03-08]. <http://cdls.nstl.gov.cn/cdls2/w3c/2003/Process/>
- [4] 孙一钢, 龙伟. 我国数字图书馆标准规范建设: 数字资源加工操作指南研究报告. [EB/OL]. [2014-03-08]. <http://cdls.nstl.gov.cn/cdls2/w3c/2003/Process/>
- [5] 姜秀华等. 数字电视广播原理与应用[M]. 北京: 人民邮电出版社, 2007
- [6] ITU-T 建议 G.1080(2008) IPTV 服务的体验质量要求 (Quality of experience requirements for IPTV services)
- [7] National Initiative For A Networked Cultural Heritage. The NINCH guide to good practice in the digital representation and management of cultural heritage materials. [EB/OL]. [2014-03-08]. <http://www.nyu.edu/its/humanities/ninchguide/>
- [8] Carignan Y, Evander J, Gueguen G, et al. Best practice guidelines for digital collections at University of Maryland Libraries[J]. Office of Digital Collections and Research, University of Maryland, College Park, MD., Second Edition Draft released April, 2007, 17
- [9] Arms C, Fleischhauer C, Jones J. Sustainability of digital formats planning for the Library of Congress collections[J]. Retrieved January, 2005, 24: 2006
- [10] Library of Congress. Digital Audio-Visual Preservation

Prototyping Projects: Appendix5: Special Consideration for Digital Video and Audio. [EB/OL]. [2014-03-08]. <http://www.loc.gov/rr/mopic/avprot/rfq5.html>

[11] MeWiki. X264 Settings. [EB/OL]. [2014-03-08]. <http://mewiki.project357.com/wiki>

[12] Santa-Cruz D, Ebrahimi T, Askelof J, et al. JPEG 2000 still image coding versus other standards[C]//International Symposium on Optical Science and Technology. International Society for Optics and Photonics, 2000: 446-454