Refining Conversion Contract Specifications: Determining Suitable Digital Video Formats for Medium-term Storage.

George Blood, principal author*

Table of contents

What is this document?

This paper was drafted by George Blood Audio and Video for the Library of Congress Office of Strategic Initiatives (OSI). OSI manages the Library's main external contract (LCOSI10C0099) for the reformatting of paper documents, still photographs, and audiovisual materials. This contract is held by the Crowley Company (http://www.thecrowleycompany.com/) of Frederick, Maryland, and George Blood Audio and Video (http://www.georgeblood.com/) is the audiovisual specialist subcontractor.

* The author wishes to express his appreciation to the many people who contributed to this paper. Carl Fleischhauer and Jimi Jones represented the Library of Congress Office of Strategic Initiatives in framing the initial assignment, refining the discussion, and helping draft the prose introduction. James Snyder, Senior Systems Administrator at the Library of Congress Packard Campus, served as the Library's expert video engineer. Snyder's input and challenges sharpened the focus of each draft and increased the accuracy of our findings. His knowledge and humor are greatly appreciated. George Blood Audio and Video staff Tim Mullin, Video Engineer, Cassandra Gallegos, Preservation Administrator, and Kimberly Peach, Registrar contributed as research assistants. Their varied perspectives contributed to the readability and usability of this document.

[1] Formats in Category 2 allow direct access to the encoded bit stream. These bits can be migrated intact in a "Tape to File" transfer, as though the information were "bit scraped" off the tape and put into a file. See details in discussions below.

Video preservation in a digital context

Video formats and carriers are inherently obsolescent. All analog video formats are obsolete[2], some severely endangered. All tape-based digital video formats are either obsolete or will be soon. File based digital video formats are evolving rapidly. The physical carriers of these files are unreliable; manufacturers assume they will be used for acquisition then transferred to hard disc drive for production. For these reasons our starting assumption is that all historic video formats must be migrated into file-based digital format.

One of the guiding principles in this paper is that all archives, regardless of format, benefit from low variability in their holdings. Ideally all holdings would be uniform. Uniformity and open standards enhance the long term preservation of digital content. All archives seek an evergreen format that will diminish the need for format migration due to obsolescence of the encoding and wrapper. This is in reach for sound recordings, where archive could, as a practical matter, produce (or migrate) all of their digital files encoded as pulse code modulated (PCM) audio, wrapped in a Broadcast WAVE file, at a limited set of resolutions, e.g., 96kHz 24 bit, 48kHz 24 or 16 bit, and 44.1kHz 16 bit. Alas, as this report makes clear, this level of uniformity is not yet in reach for video files. Over time, however, we expect both born-digital and born-analog-then-digitized formats for preservation to converge in much the way Broadcast WAVE has become the current accepted standard for audio preservation.

The scope of discussion in this paper concerns only long-term preservation master formats. Mezzanine or production-master formats, and general access and streaming formats are not addressed in this document.

Practice variation in the audiovisual preservation community

At this writing, there is considerable variation in the types of video files produced by preservation archives and, indeed, many archives are deferring the production of such files due to the uncertainties associated with emerging practices.

One important leader in the field is the Library of Congress Packard Campus of the National Audio-Visual Conservation Center in Culpeper, Virginia. The Packard Campus employs the lossless JPEG2000 encoding wrapped in the MXF file format as their long-term, "evergreen" preservation format for video. Mathematically lossless codecs are appropriate representations of uncompressed video in a preservation context. While other mathematically lossless codecs exist, none are more attractive than JPEG2000 with regards to interoperability or wide adoption.

As part of their work within the OSI contract mentioned above, George Blood Audio and Video provides the Library with deliverables consisting of files encoded with lossless

[2] "Obsolete" is defined as playback machines are no longer in production

JPEG2000 and wrapped in MXF, files that meet the Packard Campus specification. Nevertheless, OSI was interested in learning about other approaches that may have value to other archives and OSI's partner institutions, e.g., members of the National Digital Stewardship Alliance (NDSA; http://www.loc.gov/today/pr/2010/10-178.html) In part, this interest in alternate approaches reflects the emergent nature of the JPEG 2000/MXF specification.

Detailed specifications for JPEG2000 encoding conjoined with the MXF wrapper are still under development. A standardization effort is under way within the Audio-Visual Working Group of the Federal Agencies Digitization Guidelines Initiative under the rubric *Application Specification for Archiving and Preservation* (AS-AP).[3] Although the publication of the ISO broadcast profiles for JPEG2000 helped standardize this encoding for video picture information, the details for the packaging and labeling of interlaced frame images are unresolved, making JPEG2000 interoperability between different manufacturers uncertain.[4] Furthermore, common desktop applications have limited ability to access either JPEG2000 or MXF; therefore, specialized or professional tools are required to access and manage these formats.

The Library of Congress Packard Campus is a professional, well-staffed, and well-equipped organization and their circumstances have permitted them to implement the MXF/JPEG 2000 approach. They have worked with vendors to refine their specification for the best results. The Packard Campus began to use this approach in 2009 and currently has preserved more than 32,000 standard definition titles in this format. They anticipate applying the same approach to high definition video by the end of 2011 or early in 2012.

Options for an interim approach

Some archives beyond the Packard Campus, however, may wish to wait before adopting an JPEG2000/MXF approach. Some may not have a high level of professional staffing or equipment, or they may wish to wait until the AS-AP specification is more mature. These archives seek an intermediate solution, an "interim-master format" or "transitional repository format."

Since all digital information will be migrated to keep ahead of carrier obsolescence, the question comes down to *when* an specification for JPEG2000/MXF such as AS-AP can be easily implemented, permitting an archive to will catch up with any backlogs. This is our *interim* period, somewhat arbitrarily defined for this report as from three to ten years (from 2011). During this period, the selection of one or more transitional formats to be used in digital

---

[3] http://www.digitizationguidelines.gov/guidelines/MXF_app_spec.html

[4] A significant technical issue was encountered when preparing AS-AP, how to describe interlacing. MXF is built upon many existing standards. The relevant standard, SMPTE ST422, does not address interlacing. A request to open ST422 for revision to cover this omission, and thus standardize interlacing in MXF, as made in August 2011. The revision process is expected to take one year.

preservation is driven by three factors: the current format of the assets, the window of obsolescence of that format, and the point at which it is transcoded into AS-AP.

The recommendations in this report offer answers to the following questions:

- What formats will remain viable during the three to ten year period defined?
- What formats permit the retention of all of the "essential features" of the source items, i.e., formats that do not represent a loss of picture and sound quality and also retain metadata, closed captioning, or other functional features of the original?

About classes and categories

These recommendations consider two broad classes of source materials: (a) ones for which signal conversion or transcoding is necessary (e.g., analog videotapes, and obsolete file-based digital video) and (b) ones for which the native encoding can reasonably be maintained for three to ten years, after which a conversion or migration will be warranted.

Regarding class (a), examples for which signal conversion or transcoding is necessary when making the digital archival master, two general principles can be articulated. First, lossy compression is ruled out. All specialists agree that lossy "target" codecs are unacceptable for preservation, including for interim storage. Even very high bit-rate compressed formats are not acceptable due to the nature of their encoding, which causes significant degradation in transcoding (cumulative loss, concatenation errors at macroblock boundaries, etc). For this reason, uncompressed video streams are prominent in the recommendations that follow.

The second general principle honors the Packard Campus staff preference for 10-bits-per-sample resolution: digital archival masters produced for the Library must have this bit depth. The visual difference between 10-bit and much smaller 8-bit files may be subtle, especially in low grade formats such as VHS, UMatic and Betamax. Using 10 bits over 8 bits does not increase the range of information that can be captured (between 5 and 110 IRE). It does, however, increase the granularity, the level of details, between the range that is captured, with significant reduction in banding, providing smooth transitions within each chrominance or luminance channel between colors and black and white. It is worth restating that in preservation, any loss of information, especially actual instead of theoretical, is unacceptable. Low grade or low quality source material may lack in resolution, especially vertical color resolution, and may lack the full range of luminance and color. Capturing the finer detail and subtle gradations using 10-bit is therefore more important given the inherent limitations of these formats.

Regarding class (b), examples for which signal conversion or transcoding is not necessary, most specialists agree that retention of the native or acquisition format makes sense for the medium term. It is worth emphasizing however, that in some cases, when the content is more or less in a media dependent format, e.g., a DV tape, that the transfer of the bitstream without transcoding must be done in a proper manner in order to, for example, retain all embedded metadata. This non-transcoded transfer is contrasted with transcoded transfers in which, for example, the video is played out in serial streams via SDI or HDSDI and re-encoded at the point of capture.

The preceding paragraphs name two broad classes of source materials. These can be broken into five categories and these categories serve to organize the remainder of this document. The discussion that accompanies the statements of recommended delivery formats (output formats) provides an explanation of the differences.

1. Analog source
2. Digital source (media dependent, non-transcoded transfer possible)
3. Digital source (media dependent, required transcoded transfer)
4. Digital source (media independent "file based")
5. Digital source (authored DVDs and other authored disks)

Readers should note that the main part of this document concerns only preservation of the audiovisual essence. Important discussions on other content and related issues are included in Appendices. The Appendices are not meant to address these other topics comprehensively; merely to touch on the essential issues to be considered.

Other important preservation issues are beyond the scope of this document, but worth noting. It is assumed all equipment is in optimum performing condition and experienced operators are running the equipment and know how to adapt to special situations. Additionally, a comprehensive preservation environment, such as that described in OAIS, is in place, including file naming, asset management, descriptive-, technical-, administrative- and other metadata, means of providing user access, a data storage strategy covering authentication and data integrity (multiple copies, checksums, etc.), and quality control systems, both human and automated.

It is worth clarifying a few of the conventions used throughout the document.

- “Native” is used to describe a parameter as it exists on the source media. It is “same as original”.
- This document is not intended to be exhaustive, only to cover the overwhelming majority of content most collections are likely to encounter. Custom, experimental and extremely rare implementations are not addressed.
- Color space discussion is limited. Encoding should not change between the original and its surrogate: it should remain Native.
- The terms NTSC and PAL refer to the specifics of encoding their respective analog color subcarrier frequencies. In the digital domain these attributes are not present. For this reason we refer to the number of vertical lines, 525 and 625, to describe the parallel digital

systems. Another way of thinking about this would be say “Adherents to North American Systems” and “Adherents to European Systems”; or for parallelism throughout the document: NTSC/525, and PAL/625.

Category 1. Analog Source

Recommended delivery specification, NTSC/525 line

Method: Playback into encoder, wrap output as file

Wrapper: .mov (QuickTime) or .avi[5]

Video:

Video Compression: Uncompressed
Frame Size: Standard Definition: 720x486[6]
Frame Size: High Definition: Native
Aspect Ratio: Native: 4:3 for SD; 16:9 for HD
Bit Depth: 10-bit
Color Space: YCbCr
Chroma subsampling: 4:2:2
Interlaced/Progressive: Native
Frame Rate: Native, 30 or 29.97

Time Code: Native if present; Midnight start & NDF if synthetic

Audio:

Audio channels: same as original[7]
Audio Compression: Uncompressed, PCM
Audio Sample Rate: 48 kHz
Audio Resolution: 24-bit

Obsolescence Monitoring Comment: Highly stable, probably won't be necessary to transcode in 3-10 years before JPEG2000/MXF migration.

Considerations for Access: Current model computers should be able to decode this data stream, though a codec plug-in may be necessary. Playback over a computer network, however, may be impaired. Internet playback is not possible at this time.

Discussion

In order to prepare for eventual migration into JPEG2000 form, content in this category must be captured as uncompressed. Ideally all born-analog formats should be digitized as 10-bit uncompressed. At some future migration, these files will be converted to lossless JPEG2000/MXF.

---

5 Wrappers are discussed in Appendix A.

6 See Appendix B for additional discussion of resolution, raster, and aspect ratio.

7 If time code is present in an audio channel there are two schools of thought. Strict preservationists advocate for the audio/time code to be digitized as is, as well as decoded and captured as time code. Others advocate for following original intent, capturing the time code as time code and avoiding the annoyance to the user of hearing time code in one channel. The preservationists respond that failing to capture the audio channel with time code alters the representation of the original artifact and thereby removes the provenance and original organization; after all the end user can always just turn off the annoying channel.

Category 1. Analog Source

Recommended delivery specification, PAL/625 line

Method: Playback into encoder, wrap output as file

Wrapper: .mov (QuickTime) or .avi[8]

Video:

Video Compression: Uncompressed
Frame Size: Standard definition: 720 x 576
Frame Size: High Definition: Native
Aspect Ratio: Native: 4:3 for SD; 16:9 for HD
Bit Depth: 10-bit
Color Space: YUV
Chroma subsampling: 4:2:2
Interlaced/Progressive: Native
Frame Rate: Native, 25

Time Code: Native if present; Midnight start & NDF if synthetic

Audio:

Audio channels: same as original[9]
Audio Compression: Uncompressed, PCM
Audio Sample Rate: 48 kHz
Audio Resolution: 24-bit

Obsolescence Monitoring Comment: Highly stable, probably won't be necessary to transcode in 3-10 years before JPEG2000/MXF migration.

Considerations for Access: Current model computers should be able to decode this data stream, though a codec plug-in may be necessary. Playback over a computer network, however, may be impaired. Internet playback is not possible at this time.

Discussion

In order to prepare for eventual migration into JPEG2000 form, content in this category must be captured as uncompressed. Ideally all born-analog formats should be digitized as 10-bit uncompressed. At some future migration, these files will be converted to lossless JPEG2000/MXF.

[8] Wrappers are discussed in Appendix A.
[9] If time code is present in an audio channel there are two schools of thought. Strict preservationists advocate for the audio/time code to be digitized as is, as well as decoded and captured as time code. Others advocate for following original intent, capturing the time code as time code and avoiding the annoyance to the user of hearing time code in one channel. The preservationists respond that failing to capture the audio channel with time code alters the representation of the original artifact and thereby removes the provenance and original organization; after all the end user can always just turn off the annoying channel.

Category 2. Digital source (media dependent, non-transcoded transfer possible),

Recommended delivery specification, NTSC/525 line

Method: Transfer data to file wrapper without transcoding

Wrapper: Native-as-associated with the underlying encoded essence, (such as .dv, .imx, .mpeg, .mp4, etc.), or .mov (QuickTime) or .avi[10]. Otherwise wrap in QuickTime.

Video:

Video Compression: Native
Frame Size: Native
Aspect Ratio: Native: 4:3 for SD; 16:9 for HD
Bit Depth: Native, 8-bit or 10-bit
Color Space: Native, YCbCr
Chroma subsampling: Native, typically, 4:2:2, 4:1:1, or 4:2:0
Interlaced/Progressive: Native
Frame Rate: Native

Time Code: Native when present; Midnight start & NDF if synthetic

Audio:

Audio channels: Same as original
Audio Compression: Native, typically uncompressed, PCM
Audio Sample Rate: Native, typically, 48 kHz
Audio Resolution: Native, typically 24- or 16-bit

Obsolescence Monitoring Comment: Stable, one transcode may be necessary in 3-10 years before JPEG2000/MXF migration.

Considerations for Access: Current model computers should be able to decode these data streams, though a codec plug-in may be necessary. Some formats are special implementations of standards, such as IMX is a variant of MPEG2. The variety of formats adds complexity to managing an archive. Playback over a computer network, however, may be impaired. Internet playback is not possible at this time.

[10] The decision to retain the original wrapper is subject to the same considerations discussed in Category 4. Wrappers are discussed in Appendix A.

Discussion

Whenever possible, tape-based digital video should be captured in its raw state from the tape, complete with associated metadata. Tools are available to transfer the data without transcoding (the equivalent to "ripping" an audio compact disc), that is, extract the 1s and 0s as formatted on the tape. This has several advantages including faster than real-time transfer, documentation of error correction activity, and the metadata embedded at the time of recording. This last may include not only time code, but also time of day and GPS location information. This occurs prior to error concealment, which would be present at the analog or SDI output. Creating "uncompressed" versions, that is, decompressing the data, creates larger files with no additional resolution. For some formats, there are no tools available at this time to conveniently convert the metadata to a standardized format. At some future migration, these files will be decompressed to 10-bit uncompressed (with this action captured in the process history metadata), the metadata harvested and formatted, and converted to lossless JPEG2000/MXF.

Examples of formats in Category 2:

DV-family (all)
BetacamSX (in some cases)
P2
XDCAM
HDV
IMX (in PD ["Professional Disc"] forms)

Category 2. Digital source (media dependent, non-transcoded transfer possible)

Recommended delivery specification, PAL/625 line

Method: Transfer data to file wrapper without transcoding

Wrapper: Native-as-associated with the underlying encoded essence, (such as .dv, .imx, .mpeg, .mp4, etc.), or .mov (QuickTime) or .avi[11]. Otherwise wrap in QuickTime.

Video:

Video Compression: Native
Frame Size: Native
Aspect Ratio: Native: 4:3 for SD; 16:9 for HD
Bit Depth: Native, 8-bit or 10-bit
Color Space: Native, YUV
Chroma subsampling: Native, typically, 4:2:2, 4:1:1, or 4:2:0
Interlaced/Progressive: Native
Frame Rate – Native

Time Code: Native when present; Midnight start & NDF if synthetic

Audio:

Audio channels: Same as original
Audio Compression: Native, typically uncompressed, PCM
Audio Sample Rate: Native, typically, 48 kHz
Audio Resolution: Native, typically 24- or 16-bit

Obsolescence Monitoring Comment: Stable, one transcode may be necessary in 3-10 years before JPEG2000/MXF migration.

Considerations for Access: Current model computers should be able to decode these data streams, though a codec plug-in may be necessary. Some formats are special implementations of standards, such as IMX is a variant of MPEG2. The variety of formats adds complexity to managing an archive. Playback over a computer network, however, may be impaired. Internet playback is not possible at this time.

[11] The decision to retain the original wrapper is subject to the same considerations discussed in Category 4. Wrappers are discussed in Appendix A.

Discussion

Whenever possible, tape-based digital video should be captured in its raw state from the tape, complete with associated metadata. Tools are available to transfer the data without transcoding (the equivalent to "ripping" an audio compact disc), that is, extract the 1s and 0s as formatted on the tape. This has several advantages including faster than real-time transfer, documentation of error correction activity, and the metadata embedded at the time of recording. This last may include not only time code, but also time of day and GPS location information. This occurs prior to error concealment, which would be present at the analog or SDI output. Creating "uncompressed" versions, that is, decompressing the data, creates larger files with no additional resolution. For some formats, there are no tools available at this time to conveniently convert the metadata to a standardized format. At some future migration, these files will be decompressed to 10-bit uncompressed (with this action captured in the process history metadata), the metadata harvested and formatted, and converted to lossless JPEG2000/MXF.

Examples of formats in Category 2:

DV-family (all)
BetacamSX (in some cases)
P2
XDCAM
HDV
IMX (in PD ["Professional Disc"] forms)

Category 3. Digital source (media dependent, transcoded transfer required)

Recommended delivery specification, NTSC/525 line

Method: Playback, capture of SDI[12] or HDSDI serial stream, wrap file

Wrapper: .mov (QuickTime) or .avi[13]

Video:

Video Compression: SDI or HDSDI output (decompressed from lossy native encoding)
Frame Size: Standard definition: 720x486[14]
Frame Size: High Definition: Native
Aspect Ratio: Native: 4:3 for SD; 16:9 for HD
Bit Depth: Native, 8-bit or 10-bit
Color Space: Native, YCbCr
Chroma subsampling: 4:2:2
Interlaced/Progressive: Native
Frame Rate: Native

Time Code: Native when present; Midnight start & NDF if synthetic

Audio:

Audio Channels: same as original
Audio Compression: Uncompressed, PCM (potentially lossy decompressed, output on SDI or HDSDI)
Audio Sample Rate: Native, typically, 48 kHz
Audio Resolution: Native, typically 24 or16-bit

Obsolescence Monitoring Comment: Stable, one transcode may be necessary in 3-10 years before JPEG2000/MXF migration.

Considerations for Access: Current model computers should be able to decode this data stream, though a codec plug-in may be necessary. Playback over a computer network, however, may be impaired. Internet playback is not possible at this time.

---

[12] There are four variants of SDI: A & B encode 4fsc composite, 525 and 625 respectively; and C & D encode component, 525 and 625 respectively. A & B 4fsc data streams are obsolete and should not be used. Better reformatting results are generally achieved by depending on the playback machine to convert the native composite to component and output SDI C or D.
[13] Wrappers are discussed in Appendix A.
[14] See Appendix B for additional discussion of resolution, raster, and aspect ratio.

Discussion

Some formats do not allow access to the digital video as stored on tape. Digital Betacam is one well-known example. In these formats, when the tape is played, the VTR decompresses the data from the tape, and makes 10-bit uncompressed video and audio data available through an SDI (serial digital interface) on the machine. If digital video data is not available, the output of SDI should be captured as 10-bit uncompressed files. At some future migration, these files will be converted to lossless JPEG2000/MXF.

Examples of formats in Category 3[15]:

D-1
D-2
D-3
Digital Betacam
D-5
D-6
D-9
HDCAM/HDCAM SR
BetacamSX (more likely)
IMX (in tape-based form)

[15] See Footnote 9, above

Category 3. Digital source (media dependent, transcoded transfer required)

Recommended delivery specification, PAL/625 line

Method: Playback, capture of SDI[16] or HDSDI serial stream, wrap file

Wrapper: .mov (QuickTime) or .avi[17]

Video:

Video Compression: SDI or HDSDI output (decompressed from lossy native encoding)
Frame Size: Standard definition: 720x576[18]
Frame Size: High Definition: Native
Aspect Ratio: Native: 4:3 for SD; 16:9 for HD
Bit Depth: Native, 8-bit or 10-bit
Color Space: Native, YCbCr
Chroma subsampling: 4:2:2
Interlaced/Progressive: Native
Frame Rate: Native

Time Code: Native when present; Midnight start & NDF if synthetic

Audio:

Audio Channels: same as original
Audio Compression: Uncompressed, PCM (potentially lossy decompressed, output on SDI or HDSDI)
Audio Sample Rate: Native, typically, 48 kHz
Audio Resolution: Native, typically 24 or16-bit

Obsolescence Monitoring Comment: Stable, one transcode may be necessary in 3-10 years before JPEG2000/MXF migration.

Considerations for Access: Current model computers should be able to decode this data stream, though a codec plug-in may be necessary. Playback over a computer network, however, may be impaired. Internet playback is not possible at this time.

---

[16] There are four variants of SDI: A & B encode 4fsc composite, 525 and 625 respectively; and C & D encode component, 525 and 625 respectively. A & B 4fsc data streams are obsolete and should not be used. Better reformatting results are generally achieved by depending on the playback machine to convert the native composite to component and output SDI C or D.
[17] Wrappers are discussed in Appendix A.
[18] See Appendix B for additional discussion of resolution, raster, and aspect ratio.

Discussion

Some formats do not allow access to the digital video as stored on tape. Digital Betacam is one well-known example. In these formats, when the tape is played, the VTR decompresses the data from the tape, and makes 10-bit uncompressed video and audio data available through an SDI (serial digital interface) on the machine. If digital video data is not available, the output of SDI should be captured as 10-bit uncompressed files. At some future migration, these files will be converted to lossless JPEG2000/MXF.

Examples of formats in Category 3[19]:

D-1
D-2
D-3
Digital Betacam
D-5
D-6
D-9
HDCAM/HDCAM SR
BetacamSX (more likely)
IMX (in tape-based form)

[19] See Footnote 9, above

Category 4. Digital source (media independent "file based")

Recommended delivery specification, NTSC/525 line

Method: Migrate file based video to a common, preservation-appropriate format, such as migrating from flash-RAM cards, mobile phones, optical storage (DVD-ROM, BluRayROM, Professional Disc, CD-V, etc).

Wrapper: Native (.dv, .imx, .mpeg, .mp4, etc.), or .mov (QuickTime) or .avi[20]

Video:

Video Compression: Native
Frame Size: Standard definition: 720x486[21]
Frame Size: High Definition: Native
Aspect Ratio: Native: 4:3 for SD; 16:9 for HD
Bit Depth: Native, 8-bit or 10-bit
Color Space: YCbCr
Chroma subsampling: Native, typically, 4:2:2, 4:1:1, or 4:2:0
Interlaced/Progressive: Native
Frame Rate: Native

Time Code: Native when present; Midnight start & NDF if synthetic

Audio:

Audio channels: same as original
Audio Compression: Native, typically MP3, MP4, AAC, Dolby Digital or PCM
Audio Sample Rate: Native, typically 48 kHz
Audio Resolution: Native, typically 16-bit

Obsolescence Monitoring Comment: Moderately stable, one transcode may be necessary in 3-10 years before JPEG2000/MXF migration. However, a complete statement regarding file format obsolescence monitoring for these sources and recommendations for transcoding is broad and beyond the scope of this document. Nonetheless, if a digital file format is deemed obsolete or otherwise unsustainable, at the time it is to be transferred into the Library's collection it should be lossy-decompressed to 8- or 10-bit (based on the original) uncompressed.

[20] Wrappers are discussed in Appendix A.
[21] See Appendix B for additional discussion of resolution, raster, and aspect ratio.

Considerations for Access: Current model computers should be able to decode these data streams, though a codec plug-in may be necessary. Some formats are special implementations of standards, such as IMX is a variant of MPEG2. The variety of formats adds complexity to managing an archive. Playback over a computer network, however, may be impaired. Internet playback is not possible at this time.

Discussion

File based digital video exists in a wide range of formats. Some of these formats (such as P2) are complex and have dependencies, the loss of which renders the entire file difficult, if not impossible to recover. Physical carrier obsolescence is assumed. Each of these formats should be evaluated for transparency and obsolescence at the point of acquisition. They may be stored in their native-born format if not obsolete or fragile[22]. If they are unstable or obsolete they should be lossy-decompressed to 10-bit, metadata harvested and formatted, and at some future migration converted to lossless AS-AP.

[22] File format obsolescence is a moving target; this document is primarily focused on migrating from "tangible-media" based video.

Category 4. Digital source (media independent "file based")

Recommended delivery specification, PAL/625 line

Method: Migrate file based video to a common, preservation-appropriate format, such as migrating from flash-RAM cards, mobile phones, optical storage (DVD-ROM, BluRayROM, Professional Disc, CD-V, etc).

Wrapper: Native (.dv, .imx, .mpeg, .mp4, etc.), or .mov (QuickTime) or .avi[23]

Video:

Video Compression: Native
Frame Size: Standard definition: 720x576[24]
Frame Size: High Definition: Native
Aspect Ratio: Native: 4:3 for SD; 16:9 for HD
Bit Depth: Native, 8-bit or 10-bit
Color Space: YUV
Chroma subsampling: Native, typically, 4:2:2, 4:1:1, or 4:2:0
Interlaced/Progressive: Native
Frame Rate: Native

Time Code: Native when present; Midnight start & NDF if synthetic

Audio:

Audio channels: same as original
Audio Compression: Native, typically MP3, MP4, AAC, Dolby Digital or PCM
Audio Sample Rate: Native, typically 48 kHz
Audio Resolution: Native, typically 16-bit

Obsolescence Monitoring Comment: Moderately stable, one transcode may be necessary in 3-10 years before JPEG2000/MXF migration. However, a complete statement regarding file format obsolescence monitoring for these sources and recommendations for transcoding is broad and beyond the scope of this document. Nonetheless, if a digital file format is deemed obsolete or otherwise unsustainable, at the time it is to be transferred into the Library's collection it should be lossy-decompressed to 8- or 10-bit (based on the original) uncompressed.

[23] Wrappers are discussed in Appendix A.
[24] See Appendix B for additional discussion of resolution, raster, and aspect ratio.

Considerations for Access: Current model computers should be able to decode these data streams, though a codec plug-in may be necessary. Some formats are special implementations of standards, such as IMX is a variant of MPEG2. The variety of formats adds complexity to managing an archive. Playback over a computer network, however, may be impaired. Internet playback is not possible at this time.

Discussion

File based digital video exists in a wide range of formats. Some of these formats (such as P2) are complex and have dependencies, the loss of which renders the entire file difficult, if not impossible to recover. Physical carrier obsolescence is assumed. Each of these formats should be evaluated for transparency and obsolescence at the point of acquisition. They may be stored in their native-born format if not obsolete or fragile[25]. If they are unstable or obsolete they should be lossy-decompressed to 10-bit, metadata harvested and formatted, and at some future migration converted to lossless AS-AP.

[25] File format obsolescence is a moving target; this document is primarily focused on migrating from "tangible-media" based video.

Category 5. Digital source (authored DVDs, BluRay)

Recommended delivery specification, NTSC/525 line

Method: extract ISO Image of disc

Wrapper: ISO Image native (.img)

Video:

Video Compression: Native, DVD/MPEG-2, BluRay/MPEG-2 or -4
Frame Size: Standard Definition: 720x480[26]
Frame Size: High Definition: Native
Aspect Ratio: Native: 4:3 for SD; 16:9 for HD
Bit Depth: Native, 8-bit
Color Space: YCbCr
Chroma subsampling: Native, 4:2:2
Interlaced/Progressive: Native
Frame Rate: 25, 29.97 NDF, or Native (no-inverse telecine)[27]
Frame Rate: High Definition: Native

Time Code: Midnight start

Audio:

Audio channels: same as original
Audio Compression: Native, typically Dolby Digital
Audio Sample Rate: Native, typically, 48 kHz
Audio Resolution: Native, typically 16-bit

Obsolescence Monitoring Comment: Moderately stable, one transcode may be necessary in 3-10 years before JPEG2000/MXF migration. Monitor carefully since ISO Image may become obsolescent in the near- to medium-term.

Considerations for Access: Current model computers should be able to decode these data streams, though a codec plug-in may be necessary. ISO disc images are not playable with all software applications. The variety of formats adds complexity to managing an archive. Playback over a computer network should be possible. Internet playback is not possible at this time.

Discussion

See discussion below after PAL/625 specifications.

[26] See Appendix B for additional discussion of resolution, raster, and aspect ratio.
[27] On a DVD, if the source is film at 24 frames, no telecine pull-up to 30 frames is necessary. This refers to the correction applied in nominal-30-frame video production when 3:2 pulldown is used in the telecine to "stretch" the 24 film frames across 30 video frames (60 video fields) in every second. In the case of a DVD player, the pull up to correct video frame rate is done on playback, which saves valuable bits/space on the disc by moving the corrective action from the storage medium to the playback device.

Category 5. Digital source (authored DVDs, BluRay)

Recommended delivery specification, PAL/625 line

Method: extract ISO Image of disc

Wrapper: ISO Image native (.img)

Video:

Video Compression: Native, DVD/MPEG-2, BluRay/MPEG-2 or -4
Frame Size: Standard Definition: 720x480[28]
Frame Size: High Definition: Native
Aspect Ratio: Native: 4:3 for SD; 16:9 for HD
Bit Depth: Native, 8-bit
Color Space: YUV
Chroma subsampling: Native, 4:2:2
Interlaced/Progressive: Native
Frame Rate: 25, 29.97 NDF, or Native (no-inverse telecine)[29]
Frame Rate: High Definition: Native

Time Code: Midnight start

Audio:

Audio channels: same as original
Audio Compression: Native, typically Dolby Digital
Audio Sample Rate: Native, typically, 48 kHz
Audio Resolution: Native, typically 16-bit

Obsolescence Monitoring Comment: Moderately stable, one transcode may be necessary in 3-10 years before JPEG2000/MXF migration. Monitor carefully since ISO Image may become obsolescent in the near- to medium-term.

Considerations for Access: Current model computers should be able to decode these data streams, though a codec plug-in may be necessary. ISO disc images are not playable with all software applications. The variety of formats adds complexity to managing an archive. Playback over a computer network should be possible. Internet playback is not possible at this time.

---

[28] See Appendix B for additional discussion of resolution, raster, and aspect ratio.

[29] On a DVD, if the source is film at 24 frames, no telecine pull-up to 30 frames is necessary. This refers to the correction applied in nominal-30-frame video production when 3:2 pulldown is used in the telecine to "stretch" the 24 film frames across 30 video frames (60 video fields) in every second. In the case of a DVD player, the pull up to correct video frame rate is done on playback, which saves valuable bits/space on the disc by moving the corrective action from the storage medium to the playback device.

Discussion

Authored DVD-Videos may contain both video and audio, as well as subtitles, menus and interactivity. This discussion is intended to serve archives that may possess authored DVDs of such things as oral history interviews and it assumes linear representations and a single video track per disc devoid of significant interactivity.

As indicated above, we recommend that the ISO Image as found on the DVD be transferred from the disc to a server drive (or equivalent), i.e., making it media independent for management in such locations as a server environment. The ISO Image file of a DVD can be thought of as a wrapper. It is both tightly controlled by specifications ("standards based") and highly complex ("not transparent"). The enclosed video and audio formats are narrowly defined within the specification for DVD-Video, and, though highly compressed, are widely adopted. There is no immediate concern for format obsolescence of either MPEG2 video or AC3 (Dolby Digital) audio. At some point the VOB will be de-muxed and re-encoded into lossless JPEG2000/MXF.

The investigation that led us to the preceding recommendation revealed the degree to which authored DVDs are challenging to reformat. We initially studied four options: 1) capture a disc image, 2) retain the VOB, 3) de-mux the VOB and lossy-decompress the audio and video streams and re-wrap or 4) de-mux, lossy-decompress video and sound and re-wrap. The following notes provide some information about what we found.

Option 1. Capture an ISO Disc Image.

This solution retains everything possible to know about the DVD-Video and removes the information from the preservation-fragile optical media. While retaining 100% of the information on the carrier, the result is a machine-dependent object that is narrowly supported: only a narrow set of applications will play the image file.

Option 2. Retain the VOB.

This solution retains virtually everything possible to know about the DVD-Video except the main menu, and removes information from the preservation-fragile optical media. It discards the main menu that may contain a small amount of valuable descriptive metadata. It creates a digital object (.vob) that is more widely playable than an ISO Disc Image.

Option 3. De-mux the VOB, lossy-decompress audio & video streams, and re-wrap.

De-muxing creates separate files that need to be re-associated in some way, requiring operator intervention (and possible error). This option only effectively retains the video and audio essences. It discards menus and interactivity. The wrapper may be .mov, or .mpeg.[30] We note

[30] MPEG files can be stored as transport or program streams. There are subtle differences between transport and program streams. This author recommends program streams. Program streams are designed for more reliable storage and transmission (such as optical discs); transport streams are designed for less reliable transmission (such as broadcast and satellite transmission). A transport stream can hold more than one program stream. Program streams are simpler (more "transparent"), and closer in

that VOB files usually contain multiplexed MPEG-2/AC-3 in an MPEG-2 program stream. We investigated whether it might be possible to extract that program stream from the VOB file, similar to bit-scraping, or tape to files of certain types of digital videotapes. At best, we found a hit or miss result; we were unable to identify a tool capable of doing the job. This is due to a mismatch between the frame boundaries of the video data blocks and the AC-3 (Dolby Digital) data blocks. AC-3 was only intended as a delivery stream. On a disc or in digital broadcasting this is not an issue. Only when attempting to edit or put the stream on a time line (such as migrating to a file-based stream), does this become a problem. Dolby addressed this issue in their later Dolby E technology. Meanwhile, we understand that many video editors de-multiplex the disc and put the sound and video on a time line (in FinalCut or other app), and then render it out. But this approach does not produce the non-transcoded result we seek. We concluded that the creation of non-transcoded MPEG-2/AC-3 program streams from the DVD is often not possible. A mixed solution is to retain the MPEG-2 encoding and transcode the audio stream to retain synchronization. (In contrast, we note that there are plenty of tools to transcode from VOB files to other formats, which will permit the production of useful derivative files, e.g., "production masters" for such tasks as repurposing for the Web. Such production masters will support many valuable video project activities but we continue to feel that the ISO Image files copied from the DVD provide the best interim copies to support preservation goals.)

Option 4. De-mux, lossy-decompress video and audio and re-wrap.
The solution is extremely inefficient (the file size increased 30 fold with no increase in resolution) and is not recommended.

Regarding our preference (option 1 above) we understand that this recommendation to capture and retain ISO Disc Image files as an interim copy is not without risk. While many tools will read a VOB, fewer will read the ISO Disc Image. It must be noted that the DVD format overall is already waning in the marketplace and obsolescence may arrive soon. Nevertheless, retaining either the ISO Image or VOB as the transitional preservation format is practical at this time and deemed less error-prone than working from elementary streams that come from de-muxing. DVD-Video discs captured as ISO Images or VOB files must be targeted for close monitoring of format obsolescence. At each time of carrier migration, the decision needs to be made whether to de-mux and re-wrap the assets in .mov or .avi, or, if JPEG2000/MXF is sufficiently mature, to lossy-decompress the video and sound tracks, and migrate to the long-term preservation format.

---

concept to an OP1A MXF object. **However, in general there is no reason to be *making* MPEG2 streams for preservation. If you encounter an MPEG2 stream, keep it in whatever form it is, transport or program. There is no reason to change it from one to the other.**

Notes on authored High Definition BluRay DVDs

The considerations discussed for DVD-Video discs pertain to BluRay. The preservation methodology for BluRay discs and HD-DVDs would be the same as for DVDs: capture an ISO Disc Image. Keep in mind that special software or software plug-ins are required to play BluRay disc images on desktop computers. It is not possible to offer a definitive recommendation pertaining to the reformatting of BluRay authored video discs, if indeed "home made" (e.g., for oral history projects) versions of such discs exist at this time. BluRay discs contain content in EVO (Enhanced VOB) files, which carry the actual video, audio, subtitle, and menu contents in stream form. EVO files are extensions to VOB.

On a BluRay disc, the EVO files can contain video encoded in multiple formats: H.264/MPEG-4 AVC, VC-1 (Microsoft's video encoding), or MPEG-2, with the audio encoded in AC-3, E-AC-3, Dolby TrueHD, DTS, DTS-HD, PCM, and MPEG-2. At this time, there are only a few consumer software solutions that can play EVO files: VLC Player 1.2 (for Windows and Macintosh operating systems), PowerDVD, WinDVD for Windows and FFmpeg for Linux (unprotected EVO only).

In general, the best recommendation we can offer at this time is to examine any authored BluRay discs that the Library may receive, and determine whether there is merit in copying the EVO files for interim retention--a parallel approach to that recommended for standard definition DVDs--or if an alternate approach is needed. In any case, it is not possible to define a specification and develop transfer costs for this subcategory at this time.

APPENDIX A. WRAPPERS

A media wrapper is a virtual container holding the essence, a header, a block of technical metadata that describes how to play the essence, and optionally, additional information including descriptive metadata. As a rule, preservation file formats should be non-proprietary, non-vendor-specific. The MXF wrapper was developed with broad industry and user support. The specification is very broad and very detailed. As it is more widely adopted, vendor support and interoperability proven, and experience grows, it is an excellent choice in the long term. As noted in earlier sections of this document, advanced professional organizations like the Library's Packard Campus are able to implement MXF as a wrapper but this approach may be out of reach for some other archives at this time.

Our recommended non-MXF wrappers for each category are listed in the specifications in the main document. Here is a summary and a few words of explanation:

Categories for which a "new" wrapper is created at the time of reformatting:

Category 1. Analog source

Category 3. Digital source (media dependent, transcoded transfer required)

We examined the specifications for the two most widely adopted general purpose wrappers, .avi and .mov (QuickTime). These allow for additional areas and metadata, and a development specification for adding additional chunks or areas. None of these other areas are broadly adopted and interoperable industry-wide. Therefore they are considered unreliable and inappropriate for preservation. At this time the only reliable way to organize and access metadata is with an external database.

No definitive choice can be made between .avi and the QuickTime wrapper (.mov). If viewed from market dominance, or breadth of adoption, .avi is preferable as a Windows standard, but .mov is preferable because far more video production is done on Macintosh computers. It is generally possible to access either wrapper on both operating systems: tools exist to access .avi's on Macs, and .mov's on PCs. Tools exist to re-wrap the essence without transcoding.

The writers of this report have a preference for Quicktime .mov files as compared to .avi. This slight preference is based on the writers' sense of the toolset available to archives for the management of files and the preparation of access copies today, as well as to support content migration to other formats tomorrow. It is believed, however, that .avi could be used with little or no increase in risk.

Categories for which a "native" wrapper can be retained or put into play at the time of reformatting:

Category 2. Digital source (media dependent, non-transcoded transfer possible)
Category 4. Digital source (media independent "file based")
Category 5. Digital source (authored DVDs and other authored discs)

Media independent born-digital assets (Category 3) may arrive and be retained in their own "native" wrappers. Media dependent content that does not require transcoding (Category 2) may be placed in the native wrappers associated with the encoding. An example is the .dv file, a kind of "raw" data container for the DV family of video. We acknowledge, however, that a case can be made for placing essences like non-transcoded DV into .mov or .avi wrappers, and we look forward to discussions of this topic with our peers.

At this writing, our wrapper recommendations are:

- When working with media-independent born-digital assets, if the format has a wrapper already, retain the original wrapper. Otherwise, wrap in Quicktime .
- When transferring media dependent born-digital assets without transcoding, if a wrapper has been established for this encoding, use it. Otherwise, wrap in Quicktime.

The recommendation to retain or employ "native" wrappers is to more clearly declare the provenance of the asset. At the future point of migration to JPEG2000/MXF, tools should be applied to extract format-specific metadata, lossy-decompress the essence, and convert to AS-AP MXF files.

It is also worth noting that file wrappers should be monitored for obsolescence as diligently as encoding algorithms.

APPENDIX B. RESOLUTION (NUMBER OF LINES)

Vertical

A frame of NTSC analog composite video contains 525 vertical lines. The essence (video) is encoded on 486 lines. The other 39 lines, called the Vertical Blanking Interval (VBI), mostly contain synchronization signals[31]. These signals are required for analog playback, but are superfluous for digital playback. However, this space may also be used for closed captioning (line 21), and vertical interval time code (line 7 or 8). How, or even whether, to extract VBI information is outside the scope of this discussion. Nonetheless, a decision must be made about whether to look for this information and how to act upon it. See also Appendix C.

Many forms of compression organize the video frame into tiles or macroblocks, of 8x8 or 16x16 pixels. The 486 essence lines cannot be evenly divided by either 8 or 16. To simplify the encoding system, 6 lines are *often* dropped, and only 480 lines are encoded. If an encoder drops the 6 lines, they are completely lost, and will not be recreated in subsequent transcoding. A few digital video formats retain all 486.

Summary:

525: number of lines in a full frame of analog NTSC
486: number of lines containing video essence (sometimes called "active picture" or "active image"); other lines may contain metadata (closed captioning and VITC)
480: number of lines in most full-frame compressed formats

We recommend that the *masters* produced in reformatting maintain the number of lines in the original content:

- all analog sources: 486 lines*
- digital video sources with 480 lines: store 480 lines
- digital video sources with 486 lines: store 486 lines*
- digital high definition: retain what is in the source

* When these 486-line masters are used as the source to produce derivative "access" files, which will most often employ lossy compression formats, six lines will be lost from these copies in order to create the 480-line representation that fits most compression schemes.

[31] The equivalent numbers for PAL are 625 total lines, 576 lines of essence, 49 lines of VBI. For compression, 576 is divisible by 8 and 16, so the line count issue in compression is not present.

Horizontal data, square and non-square pixels

Two values are typically seen for horizontal resolution of a standard definition NTSC video source: 640 and 720 pixels wide. The aspect ratio of a standard definition NTSC frame is 4:3. If you have 486 vertical lines, to maintain "proper" aspect ratio you capture 648 horizontal pixels.

So where does this 720 number come from? There's a missing parameter. You have vertical lines, frame aspect ratio, and the missing parameter, "pixel aspect ratio." ITU-R BT.601 specifies 720 pixels, but not their shape. By increasing the horizontal resolution to 720, the narrower pixels display the proper frame aspect ratio. The pixel aspect ratio is 10:11 (width to height; that is they are narrow, or vertically tall). This yields 704 horizontal pixels. An additional 16 pixels were "padded" to the specification to allow for overscan at the edges of the frame, necessary for some images. Under many circumstances, such as display on an NTSC video monitor (as differentiated from a computer monitor), or when scaling to partial screen (such as YouTube), square pixels are more appropriate, as a 640x480 image is easy to manipulate and maintain the proper display aspect ratio. A more comprehensive discussion of this topic is available at:
http://en.wikipedia.org/wiki/Pixel_aspect_ratio

To summarize, 720x486 digitizes a full frame of standard definition NTSC video and is used for master files. 640x480 is used for derivatives.

Our recommendation: When capturing a born-digital video stream, retain its original resolution and pixel aspect ratio.

The good news is professional video playback systems accurately accommodate these differences and (largely) save the user from worrying about it. In editing and production it may be the explanation when something unusual happens, but not in playout.

APPENDIX C. SIGNAL METADATA: VBI, CLOSED CAPTIONING, ETC.

As discussed in Appendix B, the video essence is contained in 486 of 525 lines in an NTSC frame. The other 39 lines (49 lines for PAL) contain mostly synchronization signals necessary for analog playback, but superfluous to digital. This area may contain other information such as:

- Time Code (VITC): see Appendix E

- Alignment signals (equivalent to "tones and bars")
  If playback of the original tape was properly aligned at the beginning, there is no compelling reason to retain this information. Nonetheless, it is theoretically possible for the quality of the video to vary throughout the duration of the tape, and these signals along the duration may be useful to enhance playback alignment. There is no standard way to capture this information, other than to capture all 525 lines.

- CEA 608E – Closed Captioning Data on line 21 and CEA 708E – DTV Closed Captioning.

We understand that there are good, *long-term* proposals to address the retention of all of this signal metadata in wrappers like MXF. One excellent example is provided by the application specification AS-03 (MXF delivery format for PBS et al):

> Section 5.1.7 Closed Captioning. If present, CEA 608 line 21 (CC and XDS) data shall be carried in a SMPTE 334-1- and -2-2007-compliant ANC packet within a SMPTE 436M-2006-compliant VBI/ANC GC Element, using 8 bit encoding. If present, CEA 708B DTV captioning data shall be carried in a SMPTE 334-1- and -2-2007-compliant ANC packet within a SMPTE 436M-2006-compliant VBI/ANC GC Element, using 8-bit encoding. Caption language shall be specified using AMWA AS-04.
>
> Section 5.1.8 Other VBI. If present, VBI shall be carried in a SMPTE 334-1- and -2-2007-compliant ANC packet within a SMPTE 436M-2006-compliant VBI GC Element.

In addition, we understand that the AS-AP draft in progress will add a second option, one that takes advantage of the emergent practices that implement the SMPTE-TT Timed Text standards (ST 2052-0:2010, ST 2052-1:2010, and RP 2052-10:2010).

This approach cannot be used in our interim "pre-MXF" circumstance. For the Interim-Master, we identify the following possible solutions:

- Extract the closed captioning to an .srt file (.srt files are produced by the open source SubRip program that "rips" (extracts) subtitles and their timings from video).
- Consider similar alternatives to *.srt: (a) ad hoc *.vbi files, readable in some video server systems; (b) *.scc Script Files from Scenarist; and (c) *.sti files from [Spruce Tech?].
- Forward the data into the Video Ancillary Channel (VANC); this is not widely adopted.
- Capture all 525 and extract the closed caption information and embed in MXF at the time of permanent storage to JPEG2000/MXF.

We understand that, for the most part, the video likely to be sent out by the Library for contractor reformatting will *not* include closed captions, which reduces the criticality of this topic. We have no strong preference in this matter, although we see considerable merit in the widely supported "srt" approach. The SubRip file format is very basic, containing formatted plain text. The time format used is hours:minutes:seconds,milliseconds. The decimal separator used is the comma because the program was written in France. The line break used is often the CR+LF pair. Subtitles are numbered sequentially, starting at 1.

Subtitle number
Start time --> End time
Text of subtitle (one or more lines)
Blank line

This is a topic that we look forward to discussing with all interested parties.

APPENDIX D. METADATA

Textual metadata

As noted in various points in this document, metadata is an important component of this work. Metadata not only supports medium- and long-term preservation, but also the process of reformatting or otherwise transferring content from one format to another provides an excellent opportunity to capture/harvest and or record metadata about the new digital object.

The range of possible metadata is broad. From the point of view of the contract service provider, the range may be described as follows. For a given client or job, only some of this metadata will be part of the task requirement:

- Descriptive metadata. Generally speaking, this would be metadata that is received from the client in one format or another, and which the contractor may reformat to associate with the new digital objects.
- Technical metadata, file characteristics category. This is technical information about the new digital object. To a certain degree, this is automatically part of the digital target format, embedded in the data stream by the systems used to create the data. Some elements, however, may need to be added by some other process. This can be compared to many of the data elements in the AES audio metadata standard AES-57 (formerly AES-X098B).
- Technical metadata, source-item characteristics category. This is technical information about the source item. This may be provided to the contractor by the client, or may be developed by the contractor. Generally speaking, in a job with multiple objects, this data will be boilerplate for the whole batch. This can be compared to many of the data elements in the AES audio metadata standard AES-57.
- Technical metadata, process history category. This is technical information about the tools and processes used to transfer or migrate the content from the source to the new file. This can be compared to the AES audio metadata draft standard X098C, and to PREMIS ???.

How might a contractor deliver this metadata to the client? In a perfect world, there would be a standardized set of elements and structure for any or all categories of metadata and a way to embed this in delivered files. Unfortunately, such standardization does not currently exist. Even for MXF, in a preservation application, there is no guideline to turn to, either for an element list or structure or for the details for embedding. The Federal Agencies group is beginning an effort to address this shortfall but no outcome is expected in the near term.

A similar and even more severe deficit exists for file formats like QuickTime and AVI; although both of these do offer places to embed metadata, even metadata in an XML structure. For example, some commentators have suggested that a specification should be written to declare a new chunk in avi, based on the aXML chunk in broadcast wave [cf. EBU Technical

Document 3285, Supplement 5]. We disagree with this initiative for two reasons. First, the extremely limited adoption of 15 year old BWF/aXML specification, and the very fragile interoperability of broadly adopted BEXT chunk in BWF leads one to doubt the adoption of such a specification in .avi. Second, there is already an initiative for the MXF application profile for preservation. Given the time it takes to write, much less adopt specifications, it is likely MXF will be ready for use before an aXML chunk for .avi.

Metadata standards abound. When grouped as above, descriptive, technical, etc., existing standards can be applied to each sub-group of metadata. The American Library Association, Preservation and Reformatting Section (PARS) formed an ad hoc Audio Preservation Metadata Taskforce to collate existing metadata standards. Their work can be found here:

http://www.ala.org/ala/mgrps/divs/alcts/resources/preserv/audio_metadata.pdf

The PARS Standards Committee has requested a similar initiative for video. As of this writing, a taskforce has not yet been formed.

APPENDIX E. Time Code

Transfers in which time code can be inherited from the source tape.

SMPTE time code records an absolute time for every frame of video. The delivery format specifications in the main part of this document recommend that when time code data can be inherited from the source item, that it be carried over to the reformatted copy.

Time code on a video tape used as an input to reformatting can be derived from three sources:
- Control Track (CT)
- Linear Time Code (LTC)
- Vertical Interval Time Code (VITC)

All analog video formats have a reference signal (control track) to facilitate playback. This signal is very regular and has periods that correspond to the frames of video. By counting the periods of the signal, time can be derived. This is how time is displayed on consumer formats. There is no absolute reference, just "since we started counting," including the middle of the tape. It is useful due to its ubiquity.

In most circumstances time code will be contiguous from the beginning of the tape to the end of the signal. In many cases, however, it will not start at zero. For instance, the hour will be manually set and used as a tape-count reference. It is possible for the time code to change abruptly throughout the tape, such as when recording "time of day" and starting or stopping the tape while recording. Some playback devices have trouble when time code jumps suddenly. SMPTE is most-often recorded along the length of the tape. Early in its use it was recorded on an audio channel. Later formats have dedicated tracks for LTC. It is useful due to its absolute reference to a specific frame.

One limitation of LTC is that it is not possible to read when the tape is in pause. Another limitation, on some formats, is it means giving up stereo audio (recording mono audio on one channel and LTC on the other channel). To overcome these limitations, a standard was developed to put SMPTE time code in the vertical blanking interval ("The Other 39 lines" discussed in Appendix C). In an ideal world, LTC and VITC match. In the real world, they may not.

Selecting which time code source to carry over from the source tape to the reformatted copy is important, but can be challenging. The important thing to know is there may be multiple sources of time code on the same tape, and they may be different. The choice of which to use and why needs to be made on a case-by-case basis. In some cases, it may be best to abandon a jumbled time code on the source tape and apply a fresh one, as described in the first section of this appendix.[32]

Transfers in which no time code is inherited from the source tape.

The delivery format specifications in the main part of this document recommend that when time code data is *not* inherited from the source item, that the new time code use a virtual clock that is set to begin at midnight, i.e., with a setting of 00:00:00.00 (hour:minute:second:frame). In professional practice it is common to set the time code to one hour (01:00:00:00) at the beginning of the file. Indeed, under ideal circumstances 01:00:00:00 would appear at the beginning of program, after any pre-roll, black or color bars. As a practical matter this is very difficult to do, requiring careful handling after the video is captured to file.

The specifications also indicate that, when there is no inherited time code, Non Drop Frame (NDF) time code be employed. This term means that every video frame receives a number in the time code sequence. Since the actual frame rate of traditional "analog" video is 29.97 frames per second (colloquially stated as the nominal 30 frame rate), an "hour of time code" at the nominal rate is longer than an hour of wall-clock time by 3.59 seconds. In some circumstances, in order to march in step with the wall clock, operators employ "drop frame" time code but this is not recommended for the materials discussed in this report.

[32] The Library of Congress Packard Campus is pursuing a strategy of five time code streams, all stored in the MXF wrapper. LTC, VITC, frame count from zero (i.e. Midnight Start, NDF), time of day of encoding at the location of encoding (i.e. provenance of the time of migration) and time of day GMT (to create a baseline standard for all time of migration information).