The Association for Library Collections and Technical Services
Preservation and Reformatting Section

Minimum Digitization Capture Recommendations

June 2012

Submitted by:
Ian Bogus (Chair)
University of Pennsylvania
George Blood
George Blood Audio & Video
Robin L. Dale
LYRASIS
Robin Leech
Oklahoma State University
David Mathews
The Image Collective

Table of Contents

Preface

This document was created as a guideline for libraries digitizing content with the objective of producing a product that will not be re-digitized at a later point. Institutions can feel secure that if an item has been digitized at, or above, these specifications, they can depend on it to continue to be viable in the future. These guidelines only speak to the technical specifications of the digitized content itself and not to the larger issue of digitally preserving said content. In some cases, institutions may want to request a digital copy to preserve themselves further safeguarding materials by preserving them in multiple locations.

There have been numerous studies exploring the technical side of how items should be digitized and most institutions engaged in digitization have decided on specifications for themselves. It was not the authors' intention to duplicate this important work but rather to build on it. The authors reviewed past research, the practices at almost 50 organizations, and samples of digitized works to determine a recommendation of minimum specifications for sustainable digitized content.

Microfilm was originally seen as a cheap and compact alternative to print materials. Institutions could store their rapidly growing collections in a compact format that was easy to reproduce and share. The first standard for microform production was published in 1979 around the time when many academic libraries were taking action on the brittle book problem by establishing preservation programs. Before this date microforms were made on nitrate or acetate film neither of which have long shelf lives. Libraries could work together on non-duplicative microfilming projects once the standard was established. If a book was microfilmed according to the standard then another library could purchase a copy rather than re-film it. Items that were microfilmed before 1979 frequently needed to be re-filmed because of poor quality images or materials.

Digitization has followed a very similar path. It has been touted as a cheap, easily distributed format that can be stored in a compact space. At this point there is no standard but institutions are discussing how they can collaborate and share digitized content. Collaborative projects such as HathiTrust and the Internet Archive raise the question of whether an already digitized book can be a surrogate without re-digitizing it. Some institutions have struggled with the differences between "digitizing for access" and "digitizing for preservation." Ostensibly, content that is digitized for access is expected to have a relatively short shelf life, and content that is digitized for preservation should be sustainable for decades, if not longer.

Basis for Recommendations

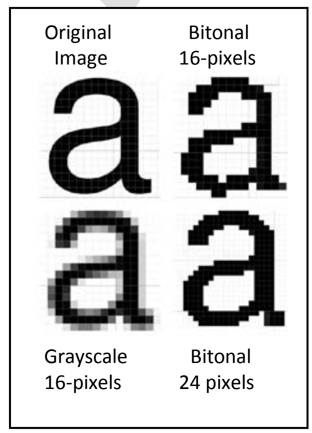
An adequate surrogate stands in place of an original by replicating characteristics that are required by users. When digitizing objects one must be aware that different types of materials are used in different ways that sometimes requires variations in the specification. The recommendations stated here should fulfill most needs, but may need to be adjusted to a higher specification in instances where the expected use is different than is described in this document. There are two general categories of media discussed below: static and time-based media. Static media is a term that encompasses common library collections such as books, photographs, maps, and microfilm that can usually be represented by image surrogates. Many resources differentiate between reflective media, where light bounces off the surface, and transmissive media, where light passes through the object. For the purposes of this document, both reflective and transmissive media are covered in the static media section. Time-based media, such as audio and video recordings, can be more complicated than static media to adequately capture. The nature of the various formats, playback mechanisms and the requirement that specific orders and timings must be captured and to be represented correctly make digitization of time-based media challenging. The section on digitizing time-based media follows the section on Static Media.

For further reference, please see a compendium of institutional guidelines in Appendix IV on page 31.

Static Media

Initially in digitization practice, many images of textual documents and line drawings used two colors: black and white, otherwise known as bitonal. This practice has fallen out of favor for various reasons. When a digital image is created it is essentially made of ordered blocks of color. When a block is created, the software needs to choose a color for each individual block. Perfect registration, where the block edges fall directly over the edges of an element, is almost impossible, and requires at least some interpretation during the digitization process. Bitonal images appear more pixelated than grayscale even when the size of each block is the same. Grayscale images actually depict the shape better because a light shade can be chosen for a block that only has a small portion of color. To combat this, many practices included increasing the resolution of bitonal image captures by 50%.

In the microfilm era, there was a system of determining legible quality called the Quality Index (QI) formula. This formula was updated for digital materials in an AIIM standard: *Resolution as it Relates to Photographic and Electronic Imaging*. Essentially, the index is based on how many pixels are required to represent the smallest significant character in the source (usually a lowercase "e" in text scenarios), which is essentially twice the QI number. Naturally, an acceptable resolution to render an object depends on the size of the characteristics. Color and grayscale images require at least 16-pixels to depict the smallest character with excellent detail resolution (QI-8). Bitonal images require 24-pixels. As a frame of reference, 400-ppi[1] will capture a 1mm artifact with 16-pixels and 600-ppi will capture that same artifact with 24-pixels. This means that 400-ppi is sufficient to adequately capture a document with 1mm characters (or a 6 pt font) in grayscale or color, and 600-ppi is sufficient to adequately capture the same document in bitonal. The figure to the right depicts a representation of the QI and differences between bitonal and grayscale representations.



One must weigh the increase in file size with the advantage of additional resolution. Details become sharper at increased resolution when blown up significantly. This can be helpful, but there comes a point where increasing the resolution does not have any gains. Without perfect focus, very high resolution can be lost in the blur. The figures[2] in the sections below show differences in resolutions between 200-700-ppi both at the original size and blown up nine times. They are formatted in a specific manner in order to show the differences at different resolutions. At the bottom are six cropped one inch by one inch images from a sample. The top row is 200-400-ppi and the bottom row is 500-700-ppi respectively.

[1] Pixels per inch. There is frequently confusion between PPI and DPI (dot per inch). A pixel is a representation on a computer screen; this pixel can be of any color. A dot references how something is printed. A printer has a limited amount of colors (usually three plus black) that need to be mixed in different ratios to create other colors. Since this document discusses capturing images and not printing them, it uses PPI.

[2] The University of Pennsylvania provided the images used in the examples. A special thanks to Chris Lippa for having the materials photographed to the specifications required to make the examples.

Above these tiles is the same area blown up nine time with the different strips at resolutions starting at 200-ppi on the left and increasing to 700-ppi on the right. Most images were not originally intended to be viewed under magnification. Extreme magnification can sometimes be helpful for special collection materials in limited cases. Generally, these materials should be available physically and extreme magnification is usually more helpful on an actual object than digitally.

A good gauge of line detail reference is a twenty dollar bill which is adequately resolved at 400-ppi. If an object has elements finer lines than the currency engravings that need to be represented then the resolution needs to be higher than 400-ppi. Conversely, an object with artifacts significantly larger than currency engravings may be able to be digitized at a lower resolution.

The use of targets can be very helpful when imaging objects as they allow staff to evaluate the images easily and objectively. The resolution value itself may not always be adequate because there can be size differences, lack of focus, interpolation, and other hidden issues. Targets usually have several components such as rulers, to ensure that the size of the image is captured correctly, as well as tools to evaluate image quality factors such as resolution, noise, artifacts and tone/color reproduction.

For further reading on digitizing static collections:

Association for Information and Image Management, “Resolution as it Relates to Photographic and Electronic Imaging,” Technical Report for Information and Image Management 26-1993 (1993): 18.

"JISC Digital Media." Still Images: creating new digital images. Joint Information System Committee, 2012. Web. 24 May 2012. <http://www.jiscdigitalmedia.ac.uk/stillimages/docs/category/creating-new-digital-media/>.

Kenney, Anne, and Stephen Chapman. *Tutorial Digital Resolution Requirements for Replacing Text-Based Material: methods for benchmarking image quality*. Washington DC: Commission on Preservation and Access, 1995. Web. <http://www.clir.org/pubs/reports/pub53/pub53.pdf>.

Peterson, Kit. *Standards Related to Digital Imaging of Pictorial Materials*. Library of Congress, Sep 2004. Web. 24 May 2012. <http://www.loc.gov/rr/print/tp/DigitizationStandardsPictorial.pdf>.

Specifications and metrics for Converted Content – a functional solution of the Future Digital System (FDsys). Washington DC: U.S. Government Printing Office, 2006. Web. <http://www.fdlp.gov/home/repository/doc_download/821-gpos-digitization-specification-33-final>.

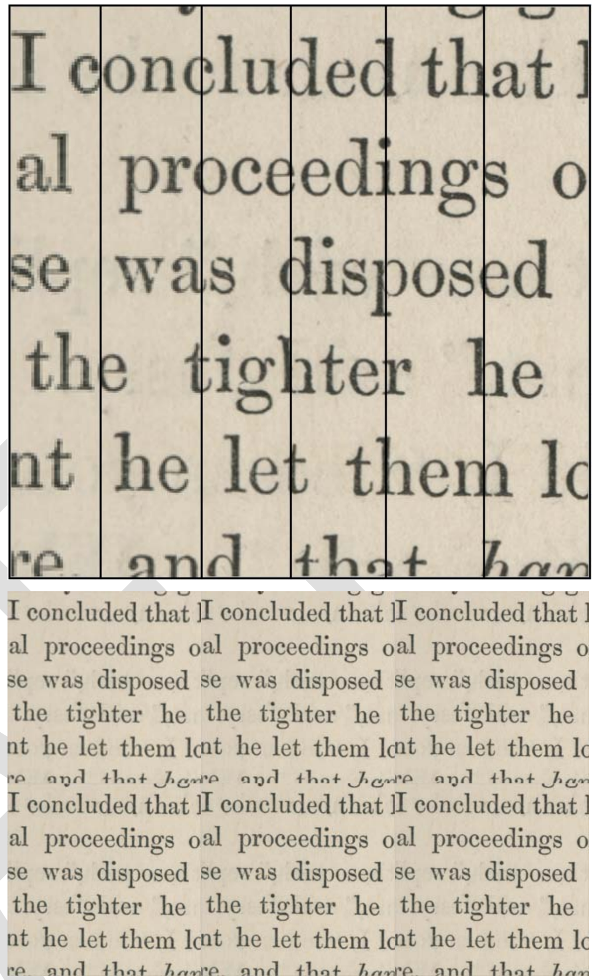
"Still Image Working Group." *Federal Agencies Digitization Guidelines Initiative*. FADGI, 27 Mar 2012. Web. 8 Jun 2012. <http://www.digitizationguidelines.gov/still-image/>.

"Universal Photographic Digital Image Guidelines." N.p., 22 Sep 2008. Web. 24 May 2012. <http://www.updig.org/index.html>.

Books and Textual Based Materials Without Images (non-rare)

The research value in most textual based materials is in the content itself. These images must be easily legible and processed through OCR or other mining software. Determining acceptable resolution depends on the size of the characters. Text larger than 1.4mm is represented by the requisite 16-pixels for excellent detail representation with 300-ppi or more. Objects that have smaller text should be imaged at a higher resolution. Grayscale images should be sufficient, but color is becoming more common and should be used whenever possible. The small “e”s in the example are approximately 1.7mm, a similar size as common 11 pt fonts..

For more information, see the suggested reading on page 5.



Above: Image blown up 9x increasing resolution from 200-ppi on the left to 700-ppi on the right

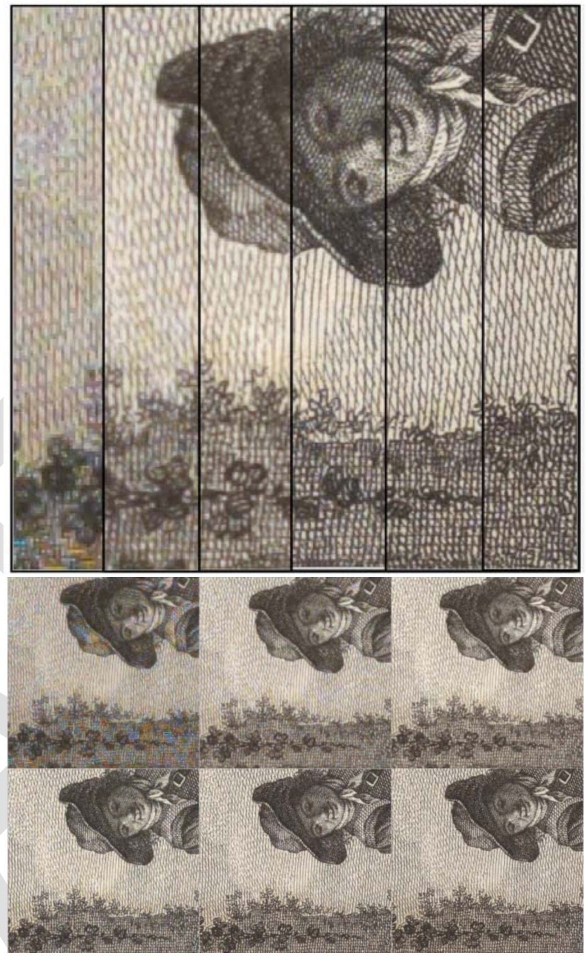
Below: Image at original size. Top row 200-400 ppi. Bottom row 500-700 ppi

Books and Textual Based Documents With Images (non-rare)

Images that are contained within books and textual based documents usually have details that can be important to users. The primary requirement is that the image is clearly. Some users are interested in the image production method, requiring clear visibility of the individual lines or dots that comprise the image that can be very small.

In most cases, 400-ppi will adequately capture necessary details. As a reference, the hash marks in the background behind the man in example are approximately 0.25mm apart. While they can be seen even at 200-ppi, increasing the resolution allows greater clarity. One can also see color artifacts at the bottom of the lower resolution images. Grayscale is sufficient for most images, but color can be helpful and should be used whenever possible.

For more information, see the suggested reading on page 5.



Above: Image blown up 9x increasing resolution from 200-ppi on the left to 700-ppi on the right

Below: Image at original size. Top row 200-400 ppi. Bottom row 500-700 ppi

Rare Books

Several factors influence methods used to digitize rare books. Less standardized fonts, or highly ornate and irregular fonts make rare books potentially more interesting to scholars; while at the same time, making legibility more difficult. Colors, stains, holes and other markings may be important to the contextual information and should be represented accurately.

Most rare books will be captured adequately at 400-ppi in color. This should capture serifs and other embellishments for any detail larger than 1mm. Contextual evidence in the paper, inks, and illustrations may need to be captured at a higher resolution. The example on has relatively large lettering measuring about 2.6mm on the small “e.”

For more information, see the suggested reading on page 5.

Above: Image blown up 9x increasing resolution from 200-ppi on the left to 700-ppi on the right

Below: Image at original size. Top row 200-400 ppi. Bottom row 500-700 ppi

Manuscripts

Manuscript materials may be written by hand and difficult to read. Additionally, there may be informational value in the inks, pen strokes, or even in the base media. Colors, stains, holes and other markings may be important to the contextual information and should be represented accurately.

Legibility and the most relevant physical information should be easily visible with 400-ppi color images. In the example, the handwriting is large with the small "e" measuring approximately 3mm. Even at this scale, faded ink is noticeably less legible at lower resolutions. In some cases, the resolution may need to be increased for legibility or when extreme magnification is necessary.

For more information, see the suggested reading on page 5.

Above: Image blown up 9x increasing resolution from 200-ppi on the left to 700-ppi on the right

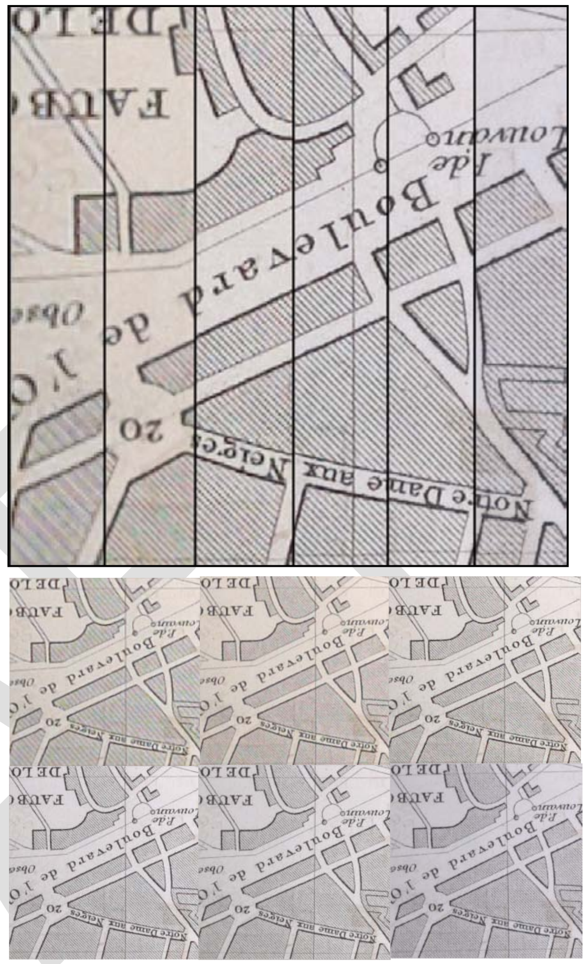
Below: Image at original size. Top row 200-400 ppi. Bottom row 500-700 ppi

Maps

Maps are quite diverse, which makes a specific digitization recommendation difficult. It is common that smaller maps have smaller details, but large maps can often have tiny elements. A digital surrogate needs to represent the original adequately for its intended purposes. Digitizing large maps at high resolutions may create exceptionally large files that are difficult to handle. Maps may also be printed at a higher print resolution than other materials.

Maps with large details may be adequately digitized at 300-ppi, though maps with very small details may require 600-ppi resolution or higher. Large file sizes may be worth the extra detail clarity for some items. The example at the right comes from a relatively small map. The “e”s in “Notre Dame aux Neiges” are all approximately 0.5 mm and the hash marks are approximately 0.25mm apart.

For more information, see the suggested reading on page 5.



Above: Image blown up 9x increasing resolution from 200-ppi on the left to 700-ppi on the right

Below: Image at original size. Top row 200-400 ppi. Bottom row 500-700 ppi

Photographic Processes
Accurately reformatting historic photographs is among the most challenging of the static media types. Prints and film are both representational media that can have subtleties in tones and colors. One should consider the intent for the digital object when selecting an adequate resolution. Digitizing for informational content is somewhat different than creating a surrogate for artifactual reasons where granules are clearly seen. Capturing the image such that the important elements are represented is usually adequate when one is concerned with the informational content, though it should be noted that photographs are commonly enlarged or magnified to view smaller elements clearly. It may be necessary to capture individual light sensitive granules when specific information on the original photographic process is important.

Digitization recommendations are commonly based on the size of the original, intending to only represent the image. Because prints and film come in different sizes, and to keep images a reasonable file size, most film and prints are grouped in three sizes with 4,000, 6,000, and 8,000 pixels along the long edge for small, medium, and large items (see the grid in the next section for details). Aerial film is an exception because the grain is much smaller and the artifacts that need to be represented are also smaller, digitized with 6,000, 8,000, and 10,000 pixels along the long edge. Aerial prints have similar grain as photographic prints, and can be captured similarly, though sometimes the discernible artifacts are smaller and may require higher resolution. There are situations where fine details are not captured adequately at the recommended minimum and resolution should be increased accordingly. True black and white images can be captured in grayscale, but color images are preferred for many photographic processes. Aerial photography is usually in black and white, but occasionally use infrared or false UV film. These latter types should be imaged in color.

Many suggest that digitizing fine art and photographic objects require greater depth and should be imaged at 16-bit grayscale or 48-bit color. There is some evidence that capturing at this higher bit depth and reducing to 8-bit grayscale or 24-bit color provides better images than imaging directly to 8-bit grayscale or 24-bit color.

Capturing granular details of the photographic process is more difficult. This is not about the minimum requirements to adequately capture an object but rather closer to the maximum level an item should be digitized. Once the granules in a photograph are fully captured, there is essentially no more information that to be captured. Photographic prints alone constitute a wide array of surface emulsions from salt prints to albumen, cyanotypes, silver gelatin, printing out paper, carbon, Platinum/Palladium, Photogravure, Bromoil, C-prints and more. Photographs have been produced for over one hundred eighty years with an expansive variance in appearance. Accurately reformatting historic photographs is one of the most challenging digitization processes of Static Media types. The majority of B&W photographs populating photographic archives are silver gelatin prints stemming from the late 1890's to present day. Color printing introduced dye layers creating additional considerations for reformatting. Investigations by the Library of Congress found that granular detail was lost when digitizing historic photographs at less than 1,200-ppi. Not all photographic prints require as high a resolution but one should try to avoid undersampling. Most historic photographic prints generically scanned result with unwanted artifacts due to uninformed workflows typified by flatbed scanning.

Photographic film scanning and post processing is technically challenging and analogous to printing in a darkroom that determines tonal values. High reformatting resolution, proper shadow and highlight values, color encoding and varied substrate issues are critical to proper workflows. Kodachrome is uniquely different than Extachrome, color film from B&W, glass plates and nitrate, Autochromes to Lantern Slides, all exhibit unique considerations.

Recent tests have shown that granules in black and white negatives from the first half of the 20th century are fully captured by 1,200-ppi but those in the second half of the century have smaller granules and need to be digitized at as high as 2,800-ppi before no further information can be captured. Some early color processes have much larger granules and captures higher than 750-ppi does not provide any more information.

Aerial Films are designed with extremely fine emulsion resolutions for recording precise subject details. Reformatting requires high resolution to preserve all subject information. The Library of Congress conducted a scientific resolution study by analyzing several aerial film types. The samplings resulted in determining the upper thresholds for reformatting resolutions. Information was discernible upwards of 2,200ppi or greater for the sample films.

For more information, see the suggested reading on page 5 and specifically for transmissive content: Williams, Don, Michael Stelmach, and Steven Puglia. *Establishing Spatial Resolution Requirements for Digitizing Transmissive Content: A Use Case Approach*. Williamson: Image Science Associates, 2011. <http://www.imagescienceassociates.com/mm5/pubs/TransmissiveResolution_Williams.pdf>.

Left: Photographic Print, Right: Photographic Film

Above: Image blown up 9x increasing resolution from 200-ppi on the left to 700-ppi on the right
Below: Image at original size. Top row 200-400 ppi. Bottom row 500-700 ppi

Posters/Broadsides/Oversize Documents

After finding equipment large enough to digitize them, the biggest hurdle with large documents is the file sizes can become large and hard to manage. Most posters, broadsides, and oversize documents are meant to be viewed from a distance and therefore do not have smaller informational elements, though one may need to print from a digital file for exhibition, reproduction, or other purposes.

Most poster, broadsides, and oversize documents will be adequately digitized at 300-ppi in 24-bit color or 8-bit grayscale depending on whether the original has colors or shades that should be represented. This should allow for a quality print reproduction, of course some documents may require higher resolution. Moiré patterning may present problems, which can be overcome in the digitization process by changing the angle of capture, software, resolution or a combination of all three.

For more information, see the suggested reading on page 5.

Art on Paper

Art on paper is a broad category of materials covering many printing, drawing, and illustration techniques. One must be particularly careful selecting appropriate resolutions for the media and intended purpose of the digital product. Many researchers are interested in the production methods of these works so magnification is not uncommon.

Works of art on paper should not be imaged below 400-ppi, 24-bit color, but there are many instances where this will be inadequate and the resolution should be increased. Moiré patterning may present problems, which can be overcome in the digitization process by changing the angle of capture, software, resolution or a combination of all three.

Microforms

Microforms are essentially photographic film with highly reduced images of textually based or manuscript materials. The film itself has a very fine grain allowing for extraordinary amount of detail in a small space. There can be great variation in the quality of microfilm depending on how it was filmed and the type of film used. In 1979 the first microfilm standard was published helping improve the quality overall.

The digitization resolution should be calculated by the size of the original document that was photographed, not the film itself. Microforms created in accordance with preservation standards will state the reduction ratio on a frame at the beginning of the roll. Because digitizing from microfilm is not imaged directly from an original, there is some concern of ill-defined letters becoming less legible because of poor registration. Most microforms should be digitized at 400-ppi with 8-bit grayscale, which accounts for some imperfection in the image quality. Very poor quality film may be adequately digitized at 300-ppi as long as registration does not further degrade the image. Continuous tone film should be scanned at 16-bit grayscale and color film at 24-bit color.

For more information, see the suggested reading on page 5.

Three-Dimensional Objects

The intended purpose of imaging a three-dimensional object is very different from the intent of many other types of digitization. One cannot currently make an adequate reproduction from a three-dimensional object, and so the intent is not to digitize so users have a surrogate to use, but rather to give a user general information about the object. Three-dimensional objects will most likely be reimaged at a later point. In libraries, three-dimensional objects are typically photographed at the native camera resolution of 300ppi. The size of the sensor and size of the capture area determine the achieved resolution. It is not uncommon that 3D objects receive several views (potentially presented as a rotating object). Lighting and camera angle of view are essential considerations and details of specific regions may be useful.

For more information, see the suggested reading on page 5.

Time-Based Media

Audio

Throughout the history of preservation, improvements in digital conversion have made it possible to capture more information from original sources. It has also left us with the impression that "tomorrow it will be possible to do better". At some point, however, we realize that there is no additional information to be captured (or the definition of what we're capturing changes). A classic example is once we can image film grains at the atomic level will it still be necessary to increase the ppi count?

All digitization is a "garbage in, garbage out" scenario. The biggest improvements in digitization come from better reproduction of the original. The biggest bang for the buck is in the analog domain. A great camera won't make a better image of a poorly lit picture.

Audio analog to digital converters now capture frequencies beyond the range of human perception, and dynamic range at the limits of the laws of physics. Further "improvements" are likely to be aesthetic, rather than technical. That is, "we like the sound better", rather than any measurable or perceptible improvements. Indeed, since Thomas Edison recorded "Mary Had a Little Lamb" on tinfoil, all audio recording technologies have captured more information than could be reproduced with the technology available at the time. It is easy to demonstrate how much "better" a CD sounds, that is how much additional information is already captured, by changing the playback. If you up-sample to 88.2, and move the anti-aliasing filter where phase distortion byproducts are outside the range of human hearing, voila – improved sound.

In audio digitization a bit is equal to 6dB. The more bits the more dynamic range you can capture. Also called signal to noise ratio, it is a measure of the range of signals from full scale (saturation) to the smallest signal that can be resolved. Changes in the operating temperature affect how quickly the atoms, even in wires, are moving. Higher temperatures mean faster movement of electrons leading to more noise. A 24 bit system can capture 144dB of dynamic range. At room temperature (68degF) an ideal electrical circuit will have a theoretical dynamic range of about 130dB, or about 21.5 bits (130/6). Even though we cannot capture 144dBs we use 24 bits because computers think in round numbers, and here 8 bits = 1 byte, and 3 bytes = 24 bits.

"Noise Calculator." *Sengpielaudio*. N.p., n.d. Web. 8 Jun 2012.
<http://www.sengpielaudio.com/calculator-noise.htm>.

Human hearing at birth is 20kHz. According to Nyquist we need a sample rate twice the highest frequency we wish to capture. Therefore, 40kHz is adequate to capture all the frequency information in an analog source that humans are able to hear. Much higher sample rates are necessary to capture information other than maximum frequency. This is due to other parameters, which aren't as well understood, where there are less concrete measurements of the limits of human perception, specifically our sensitivity to time and phase information. For this reason the upper range of sampling has moved first from 44,100 to 48kHz to 96kHz. Beyond 96kHz additional improvements must be sought in other areas of converter design. The additional information captured is overwhelmed by other factors. Over sampling and digital filters are two such strategies to improve the quality of information captured at 96kHz output. Timing information moves from the realm of how often we sample to how regularly spaced those samples are. High end converters measure this "jitter" in pico-seconds (parts per million). While it is true we are discovering human hearing is able to perceive these minute variations in timing stability, we must remember the timing errors in reproducing our sources are measured in parts per

10,000. It is possible the speed stability of playback may improve, but remember speed instability is also captured in the original recording. We are making greater representations of flaws in the original. Even more importantly the speed variations in analog playback are highly variable and random: no two playbacks will have the same variations.

Heap on top of all this the countervailing influence of another law of physics, entropy. The rate of deterioration of our analog audio sources is greater than the rate of improvement in our digitization technology. If we wait until tomorrow our digitization will improve, but we'll be capturing more deteriorated sources.

Proper mechanical and electrical function of analog playback machines requires trained specialists. Even the most low-skilled operator, however, can be trained to do playback alignment with the right tools.

Analog calibration references are available from:

"MRL Calibration Tapes." *Magnetic Reference Laboratory*. N.p., 19 Feb 2010. Web. 8 Jun 2012. <http://home.comcast.net/~mrltapes/>.

"78 rpm Calibration Disc Set." *Audio Engineering Society(AES)*. Audio Engineering Society, 2012. Web. 8 Jun 2012. <http://www.aes.org/publications/standards/calibration.cfm>.

Performance and alignment of analog to digital converters requires very specialized tools.

Simple calibration tools and standards are under development by the Federal Agencies Digitization Guideline Initiative (FADGI) in collaboration with AudioVisual Preservation Solutions. Watch their sites for developments:

"Guidelines: Audio Digitization System Performance." *Federal Agencies Digitization Guidelines Initiative*. FADGI, 15 Mar 2011. Web. 8 Jun 2012. <http://www.digitizationguidelines.gov/guidelines/digitize-audioperf.html>.

Video

Despite its ubiquity in our world today, it's hard to grasp how monumental a leap in technology video was in 1956. The number of technological heralds overcome to build a complete system nearly from scratch is inspiring. Though early video borrowed from existing technologies, such as 15 inches per second tape speed used in high quality audio and vacuum columns for tape tension borrowed from computer tape drives, the structure and organization of the information stored in and recovered from an electronic signal is very clever.

The electronic signal that is video is organized in ways that both facilitate digitization and thwart efforts to do so efficiently. Like film, video is divided into discrete elements called frames. In the US, film is 24 frames, and video is 30 frames per second. This uses the psycho-optic phenomenon "persistence of vision" to fool us into believing we're seeing motion rather than a rapid series of discrete pictures. The image portion of these frames is further organized into 486 separate horizontal lines. These discrete elements, frames and lines, lend themselves naturally to the discrete nature of digitizing otherwise analog content.

Along the lines, however, there is high frequency information modulated ("encoded") in the electronic signal. In the analog domain this information scans an electron beam across a field of multi-colored phosphors to produce an image on a cathode ray tube (CRT). Video digitization captures 720 pixel elements along each line.

The resolution of each pixel falls within a fixed range (7.5 to 110 IRE). As bits are added to resolution, the fixed range is successively subdivided into finer and finer resolution. By comparison in audio a bit has a fixed range, 6dB, and the more bits you add the greater the dynamic range. If we have 2 bits of video resolution, we have only black and white values. As we move through 3, 4 and higher resolutions we have finer and finer gradations of gray, in addition to the extreme values of black and white. To avoid banding artifacts, 10 bits of resolution are required.

In essence, video digitization is equivalent to producing 30 TIFFs each second, with a resolution of 720 x 486.

Due the very large file sizes of uncompressed video (approx 100GB per hour), some people advocate for compressing video files. The underlying arguments are no different than we wrestled with 10 years ago for still images. Indeed, at its core, a compressed video stream is a series of JPEGs.

Compression is always bad in video. Even "visually lossless" compression is a bad thing. Compression algorithms change over time. The compromises in one compression codec that fool the eye today may be exaggerated in the codec of tomorrow, yielding visible artifacts. These artifacts will be uncorrelated to the picture making them even more visible and annoying.

The biggest problem in video compression lies in the most common strategy. After digitization into 720 x 486, the field of pixels is grouped into blocks, or tiles, 8x8 square. These are lossy compressed using a mathematical technique called DCT, discrete cosine transform. Even advanced encoders have limited knowledge of what's happening in adjacent blocks. Where the blocks meet, on all four sides, there is a seam. The vertical seam is unnatural to the organization of the picture. Even the horizontal seam, though natural to the structure of video as described above, will contain some error, which is the result of the lossy compression. When the image reconstructed on playback, decoders will soften the edges to conceal the error. Any loss of information is anathema to preservation. Lossy compression followed by softening of the boundaries between the tiles further deviates from the original.

The information lost during lossy compression, and the artifacts introduced by softening the boundaries as the images are constructed cannot be restored. In a subsequent re-encoding to the latest compression technology, these errors accumulate. What might have reasonably been "visually lossless" is always preservation inappropriate, and will have additional losses over generations.

Most born-digital video is compressed at inception. As with all born-digital objects an institution must decide if they are going to support that file format and codec or not. If they are, store the digits as they are. If the file format and codec are not going to be supported within the institution's repository, the file should be decoded/decompressed to the 720 x 486 frame size and captured as an uncompressed image stream.

Appendix V (on page 34) is a white paper commissioned by the Library of Congress on specifications for digitizing video.

Moving Image Film

Film is the last area where there are large numbers of people who believe analog duplication, that is film to film, is the proper form of preservation. Preservation on film is expensive, time consuming, creates hazardous waste and is subject to the same challenges of all analog-to-analog-to-analog duplication.

There is significant disagreement within the trade about whether digitization will ever be appropriate for film. One side argues it is only a matter of time until the resolution of digitization surpasses the amount of information that can be captured in photographic film. The other side argues digital projection will never match luminesce of light projected through film onto a screen. Digits may capture the information on film, but it will never reproduce the experience of projected film.

This topic is discussed in detail in the following document produced by the Academy of Motion Pictures Arts and Sciences.

"Digital Dilemma 2." Academy of Motion Picture Arts and Sciences, 2012. Web. 8 Jun 2012. <http://origin-www.oscars.org/science-technology/council/projects/digitaldilemma2/download.php>.

We may look back upon two recent events as watersheds in the way film is preserved. In the past year, Kodak filed for bankruptcy and an earthquake in Japan shut down the only factory where XD CAM SR tape is manufactured. The demise of Kodak removes a major player in the manufacturer and processing of film. The inability to purchase tape for $130,000 professional video machines hastened the migration toward file-based workflows. As data storage has become larger and cheaper, the economics of working on computers are fast becoming significantly less than purchasing, handling, editing and storing film.

While major motion picture studios can be expected to continue to do high end color separation, photo-chemical preservation of major films, and to care for the analog preservation work performed to date, the rapidly falling cost and increasing resolution of digitization will make it harder and harder for smaller institutions to justify film-to-film preservation.

The authors feel that there are currently too many unknowns to make a well informed recommendation on digitalizing moving image film at this time.

Recommended Minimum Capture Summary:

	Min Resolution	Min Color Space	Min Bit Depth	Notes
Printed Materials (general collections - text only)	300	Grey	8	Capture in color (24 bit) whenever possible. 300-ppi will capture details larger than 1.4 mm at a QI of 8. The resolution may be adequately adjusted according to the largest detail to be represented.
Books (general collections - w/ images)	400	Grey	8	Capture in color (24 bit) whenever possible. Fine lines in etchings or other pictorial elements require more definition than text alone. 400-ppi will capture details larger than 1 mm at a QI of 8. The resolution may be adequately adjusted according to the largest detail to be represented.
Manuscripts	400	Color	48	Illegible or difficult to read scripts may require a higher resolution.
Microforms	400	Grey	8	Resolution should be calculated @ 100% of the size of the original object. 300-ppi will suffice if the film has innately low resolution (e.g. some newspapers). Use 16-bit grey if the film is continuous tone or 24-bit color if the film is color.
Rare Books	400	Color	24	
3D Objects	300	Color	24	
Aerial Photographic Prints (<8"x10")	400	Grey	8	Use 24-bit color when appropriate. Resolution is based on the commonly used 4000 pixels per long edge commonly used for photographic digitization. Use 4000 pixels whenever possible ultimately creating a higher resolution
Aerial Photographic Prints (8"x10" - 11"x14")	400-600	Grey	8	Use 24-bit color when appropriate. Resolution is based on the commonly used 6000 pixels per long edge commonly used for photographic digitization. Use 6000 pixels whenever possible ultimately

				creating a higher resolution
Aerial Photographic Prints (>11"x14")	600	Grey	8	Use 24-bit color when appropriate. Resolution is based on the commonly used 8000 pixels per long edge commonly used for photographic digitization. Use 8000 pixels whenever possible ultimately creating a higher resolution
Aerial Photographic Film (70mm - 4"x5")	1200-2150	Grey	8	Use 24-bit color when appropriate such as infrared or false color UV. Resolution is based on the commonly used 6000 pixels per long edge commonly used for photographic digitization. Use 6000 pixels whenever possible ultimately creating a higher resolution
Aerial Photographic Film (4"x5" - 5"x7")	1200-1600	Grey	8	Use 24-bit color when appropriate such as infrared or false color UV. Resolution is based on the commonly used 8000 pixels per long edge commonly used for photographic digitization. Use 8000 pixels whenever possible ultimately creating a higher resolution
Aerial Photographic Film (> 5"x7")	1450	Grey	8	Use 24-bit color when appropriate such as infrared or false color UV. Resolution is based on the commonly used 10000 pixels per long edge commonly used for photographic digitization. Use 10000 pixels whenever possible ultimately creating a higher resolution
Art on Paper	400	Color	24	
Photographic Film (35mm-4"x5")	800-2800	Grey / Color	8 / 24	Resolution is based on the commonly used 4000 pixels per long edge commonly used for photographic digitization. Use 4000 pixels whenever possible ultimately creating a higher resolution
Photographic Film (4"x5" - 8"x10")	800-1200	Grey / Color	8 / 24	Resolution is based on the commonly used 6000 pixels per long edge commonly used for photographic digitization. Use 6000 pixels whenever possible ultimately creating a higher resolution
Photographic Film	800	Grey /	8 / 24	Resolution is based on the commonly

(>8"x10"Film (4"x5" - 8"x10""x10")		Color		used 8000 pixels per long edge commonly used for photographic digitization. Use 8000 pixels whenever possible ultimately creating a higher resolution
Photographic Prints (<8"x10")	400	Grey / Color	8 / 24	
Photographic Prints (8"x10" - 11"x14")	400-600	Grey / Color	8 / 24	Resolution is based on the commonly used 6000 pixels per long edge commonly used for photographic digitization. Use 6000 pixels whenever possible ultimately creating a higher resolution
Photographic Prints (>11"x14")	600	Grey / Color	8 / 24	Resolution is based on the commonly used 8000 pixels per long edge commonly used for photographic digitization. Use 8000 pixels whenever possible ultimately creating a higher resolution
Oversize Documents	300	Grey / Color	8 / 24	Requires special handling considerations.
Maps	300-600	Grey / Color	8 / 24	600 ppi will capture highly detailed information and good for reprinting and fine quality maps. Lower resolutions may be appropriate if detail is limited (large print etc...) or will not be printed.
Posters/Broadsides	300	Color	24	The off-set print resolution of posters and broadsides may present screening artifacts creating moiré patterns requiring adjustments.
Audio	96000		24	
Analog NTSC Video	720 x 486		8	
Digital Video Source Tape, where possible to access bits	Native		Native	
Digital Video Source Tape, where not possible to access bits	Decompressed 720 x 486		8 or 10	
Digital Video File	Native			Subject to file format obsolescence evaluation. If deemed obsolete, decompress to 10-bit native raster

				(horizontal x vertical pixel count)
Video Optical Disc	Native		Native	Reformat to ISO disc image to capture all video, all angles, all subtitle and multiple languages, and menus.

Appendix I: File Naming conventions for Digital Collections

Electronic files should be well organized and named in such a way that they are easily identifiable and accessible. The examples in this appendix center primarily on documents, but the guidelines below can easily apply or be adapted to all file formats. The guidelines are considered best practices, however, not all may be relevant to everyone or every situation. These do provide a groundwork for designing a consistent and easy to use file-naming standard.

1. Use only alpha-numeric characters for both files and folders.
   - Exceptions are dashes (-) and underscores (_)

2. Do not use special characters.
   - Examples include: / > < + =^ | \ { } [ ] # these characters are used by the operating system

3. Use a three-character file extension (i.e. ".tif" ".pdf", etc.)

4. Do not use spaces in file/folder names – use dashes or underscores instead

5. Use leading zeros
   - When file names need numbering, use zeros as placeholders. Example – a collection of 900 items should be numbered: abc001.tif, abc002.tif, etc.
   - This will facilitate proper sorting.

6. When dates are used, base the date on an international standard – ISO 8601
   - Example: yyyymmdd, 20110501 or 2001_05_01

7. Be brief! Not all systems are the same, or allow for lengthy file names. Long filenames may prohibit portability

8. File names should contain necessary descriptive information, independent of their storage location
   - Examples:
     - jones_diary_18900420_0001.tif
     - Not: 00001.tif

9. If versioning is needed, include a version numeral
   - v01, v02, etc.
   - Suggestions: for the final version, include the word "final", without a version number

Appendix II: Metadata

Hundreds of metadata standards exist within the broader cultural heritage community. A fantastic overview of these often interrelated standards is available in the new resource, Seeing Standards, a Visualization of Metadata Universe[3]. Minimally, this should be basic descriptive and technical metadata collection sufficient to allow retrieval and management of the digital copies and to provide basic contextual information for the user. Additionally, the inclusion of preservation metadata through the use of PREMIS is advisable.

Below are the applicable technical and preservation metadata standards related to these recommendations.

General Standards

A Framework of Guidance for Building Good Digital Collections. 3rd ed. Baltimore: National Information Standards Organization, 2007. Web. <http://www.niso.org/publications/rp/framework3.pdf>.[4]

"Metadata Encoding and Transmission Standard." Library of Congress, Apr 2010. Web. 11 Jun 2012. <http://www.loc.gov/standards/mets/mets-schemadocs.html>.

Technical

Still Images:

"Metadata for Images in XML Standard (MIX)." *NISO Metadata for Images in XML Schema*. Library of Congress, 18 May 2008. Web. 8 Jun 2012. <http://www.loc.gov/standards/mix/>.

NISO Z39.87 2006 (R2011) *Data Dictionary - Technical Metadata for Digital Still Images*. Bethesda: National Information Standards Organization, 2006. Web. <http://www.niso.org/apps/group_public/download.php/6502/Data Dictionary - Technical Metadata for Digital Still Images.pdf>.

Audio:

AES Standard for Audio Metadata- audio object structures for preservation and restoration. New York: Audio Engineering Society, 2011. Web. <http://www.aes.org/publications/standards/search.cfm?docID=8>.

AES Standard for Audio Metadata- core audio metadata. New York: Audio Engineering Society, 2011. Web. <http://www.aes.org/publications/standards/search.cfm?docID=85>.

"Audio Technical Metadata Extension Schema." Library of Congress, 02 May 2011. Web. 8 Jun 2012. <http://www.loc.gov/standards/amdvmd/audioMD.xsd>.

"78 rpm Calibration Disc Set." *Audio Engineering Society(AES)*. Audio Engineering Society, 2012. Web. 8 Jun 2012. <http://www.aes.org/publications/standards/calibration.cfm>.

---

[3] Riley, Jenn. *Seeing Standards: A Visualization of the Metadata Universe*. 2009-2010. <http://www.dlib.indiana.edu/~jenlrile/metadatamap/>

[4] A new version of this publication is being compiled.

Video:
"Schema Documentation for videoMD.xsd." Library of Congress, n.d. Web. 8 Jun 2012. <http://www.loc.gov/standards/amdvmd/htmldoc/videoMD.html>.

Preservation (PREMIS)
Data Dictionary Section From PREMIS Data Dictionary for Preservation Metadata. version 2.1. Washington DC: Library of Congress, 2011. Web. <http://www.loc.gov/standards/premis/v2/premis-dd-2-1.pdf>.

"PREMIS: Preservation Metadata Maintenance Activity." . Library of Congress, 1 Jun 2012. Web. 8 Jun 2012. <http://www.loc.gov/standards/premis/>.

"Schemas for PREMIS." *PREMIS: preservation metadata maintenance activity*. Library of Congress, 16 May 2012. Web. 8 Jun 2012. <http://www.loc.gov/standards/premis/schemas.html>.

Structural
Still Image:
Best Practices for Structural Metadata. New Haven: Yale University Library, 2008. Web. <http://www.library.yale.edu/dpip/bestpractices/BestPracticesForStructuralMetadata.pdf>.

Descriptive
General:
Descriptive Metadata Guidelines. Mountain View: Research Libraries Group, 2005. Web. <http://www.oclc.org/research/activities/past/rlg/culturalmaterials/RLG_desc_metadata.pdf>.

"Dublin Core Metadata Element Set, Version 1.1." *Dublin Core Metadata Initiative*. Dublin Core, 11 Oct 2010. Web. 11 Jun 2012. <http://dublincore.org/documents/dces/>.

Guidelines for the Creation of Digital Collections: best practices for descriptive metadata. Champaign: Consortium of Academic and Research Libraries in Illinois, Web. <http://www.carli.illinois.edu/comms/board/Metadata_guidelines.pdf>.

"Metadata Object Description Schema: MODS." Library of Congress, 18 Oct 2011. Web. 11 Jun 2012. <http://www.loc.gov/standards/mods/>.

Still Image:

DIG35 Specification: metadata for digital images. Boynton Beach: Digital Imaging Group, 2001. Web. <http://www.i3a.org/wp-content/uploads/docs/dig35-v1.1.pdf>.

Audio/Visual:
Access Options for Embedding Metadata in WAVE Files and Plan the Audio Metadata File Header Tool Development Project. New York: AudioVisual Preservation Solutions, 2009. Web. <http://www.digitizationguidelines.gov/audio-visual/documents/AVPS_Audio_Metadata_Overview_090612.pdf>.

Embedding Metadata in Digital Audio Files. Washington DC: FADGI, 2012. Web. <http://www.digitizationguidelines.gov/audio-visual/documents/Embed_Guideline_20120423.pdf>.

"Information Technology - MPEG-21 Multimedia Framework." Klagenfurt University, 03 Jan 2005. Web. 11 Jun 2012. <http://mpeg-21.itec.uni-klu.ac.at/cocoon/mpeg21/_mpeg21Parts.html>.

Appendix III: Storage

There are two kinds of preservationists: those who have lost data and those who will. Losing data isn't in and of itself a crisis. It should be little more than a short term inconvenience. Data can be lost due to three causes: an inability to access the information due to file format obsolescence, corruption of data and failure of the physical carrier. File format obsolescence is beyond the scope of this document. For more information on tools for checking a file's viability explore JHOVE, DROID, and PRONOM. The solution to both of the other two cases is to retrieve the data from a second copy. The first action, then of any preservation strategy is to have more than one copy. Ideally that additional copy (or those additional copies) should be geographically isolated, and be on a different storage technology. Examples of two different technologies would be hard disc drives and data tape. A system to manage the data for preservation consists of a method to detect errors due to corruption and errors due to loss, either to erasure or media failure. The next stage is to have a system that both self monitors and when loss is discovered, automatically replaces the loss.

OAIS describes such a system, including the metadata for discovery and management of the content.

Open Archival Information System Reference Model: introductory guide. Dublin: Online Computer Library Center, 2004. Web. <www.dpconline.org/docs/lavoie_OAIS.pdf>.

Reference Model for an Open Archival Information System (OAIS). PINK BOOK: draft recommended standard. Washington DC: Consultative Committee for Space Data Systems, 2009. Web. <http://public.ccsds.org/sites/cwe/rids/Lists/CCSDS 6500P11/Attachments/650x0p11.pdf>.

The most widely known implementation of the data integrity management portion of OAIS is LOCKSS. Multiple copies of data are stored in disparate locations, with each node monitoring it's own health and reporting to the other nodes. When there is loss, the other nodes provide a replacement copy of the lost or corrupted file. If a node is making frequent requests for replacements it is deemed untrustworthy, administrators are notified automatically, and corrective action is taken. LOCKSS can be installed and configured in less than an hour.

"LOCKSS: Lots of Copies Keep Stuff Safe." Stanford University, n.d. Web. 8 Jun 2012. <http://www.lockss.org/>.

Preservation systems apply "fixity" values to assure files have not changed, and thereby guarantee authenticity. MD5 and SHA-1 checksums are commonly used fixity algorithms. Any computer file, regardless of size, is fed into the algorithm. It generates a "hash" value. If a single bit, or multiple bits are changed within the file, a different hash value results. Due to a high "avalanche" or "cumulative error", the values are very different due to even a small change; that is, it's easy to tell the hash value has changed. It is theoretically possible for multiple bits to change and the same hash value to result. However, you are 50 million times more likely to be struck by lightning twice than for this to happen. It is also possible for two files to have the same hash value, as there are only 2^128 combinations. As a practical matter, the high improbability of these two cases, combined with the extremely low cost of implementation makes these fixity algorithms highly useful for preservation.

As defined by the Preservation and Reformatting Section of the American Library Association, a key tenant of digital preservation is migration.

"Definitions of Digital Preservation." Association for Library Collections and Technical Services, 24 Jun 2007. Web. 8 Jun 2012. <http://www.ala.org/alcts/resources/preserv/defdigpres0408>.

Within the window of obsolescence, typically approximately 5 years, data files are migrated to the next generation of storage technology. At the point of migration you confirm the readability of the media by the act of access the file, verify authenticity with checksums, evaluate the file format obsolescence status using file validation tools, perform authority control on file names and embedded metadata, update metadata and perform preservation file retention actions, update checksums as needed, and copy to new media. As computer processing speeds have gotten faster, only verifying and generating new checksums consume much time. Indeed most of the other actions would not be noticeably slower than a simple file copy.

If all you do is copy your master files to another storage medium, by having a second copy in storage at another location, you will have performed an important first step in digital preservation. This first step is more important than having a fully deployed OAIS environment. But it is only the first step.

By exploring the Weblinks and terms above you can find many resources to learn and implement these strategies.

Appendix IV: Institutional Guidelines:

Berkeley University
http://sunsite.berkeley.edu/Imaging/formats.html

Bibliographical Center for Research
http://www.bcr.org/cdp/best/digital-imaging-bp.pdf
http://www.bcr.org/cdp/best/digital-audio-bp.pdf

California Digital Library
http://www.cdlib.org/services/dsc/tools/docs/cdl_gdi_v2.pdf

Columbia University
http://www.columbia.edu/acis/dl/imagespec.html

Cornell University
http://www.library.cornell.edu/imls/image%20deposit%20guidelines.pdf
http://www.library.cornell.edu/preservation/tutorial/conversion/table3-1.html

Digital Library Federation
http://www.diglib.org/standards/bmarkfin.htm

Digital Library of Georgia
http://dlg.galileo.usg.edu/guide.html#04

Digital Library of the Caribbean
http://ufdc.ufl.edu/design/aggregations/dloc/training/en/Section6.pdf

Federal Agencies Digitization Guidelines Initiative
http://www.digitizationguidelines.gov/guidelines/FADGI_Still_Image-Tech_Guidelines_2010-08-24.pdf

Government Publications Office
http://www.gpoaccess.gov/legacy/registry/DigitizationSpecification3.0.pdf

Historical Voices
http://www.historicalvoices.org/papers/audio_digitization.pdf

Indiana University
http://www.dlib.indiana.edu/projects/sounddirections/papersPresent/sd_bp_07.pdf

International Association of Sound and Audiovisual Archives
http://www.iasa-web.org/downloads/publications/TC03_English.pdf

Johns Hopkins University
http://www.library.jhu.edu/collections/institutionalrepository/irpreservationpolicy.html#Standards

Library of Congress
http://memory.loc.gov/ammem/about/standardsTable1.pdf

http://www.loc.gov/rr/mopic/avprot/audioSOW.html

Memory of the Netherlands Project
http://www.kb.nl/coop/geheugen/extra/downloads/Richtlijnen%20en%20procedures%20%204.0%20english.pdf

Michigan State University
http://www.historicalvoices.org/papers/audio_digitization.pdf

National Anthropological Archives
http://voom.si.edu/anthro/imaging_standards.htm

National Archives and Records Administration
http://www.archives.gov/preservation/technical/guidelines.pdf

Library of Congress
http://memory.loc.gov/ammem/dli2/html/document.html#page
http://www.loc.gov/ndnp/pdf/NDNP_200911TechNotes.pdf

National Library of Australia
http://www.nla.gov.au/digital/capture.html

New York State Archives
http://www.archives.nysed.gov/a/records/mr_erecords_imgguides.pdf

New York University
http://www.nyu.edu/its/humanities/ninchguide/VII/
http://aic.stanford.edu/sg/emg/library/pdf/mcdonough/McDonough-EMG2004.pdf

New Zealand Archives
http://continuum.archives.govt.nz/files/file/standards/s6/s6-app5.html

NISO Framework Working Group
http://www.niso.org/publications/rp/framework3.pdf

Oxford University
http://www.odl.ox.ac.uk/papers/odlprice.pdf

Penn State University
http://www.libraries.psu.edu/psul/digipres/bestpractices.html

Rutgers University
http://rucore.libraries.rutgers.edu/collab/ref/dos_avwg_video_obj_standard.pdf
http://rucore.libraries.rutgers.edu/collab/ref/dos_repteam_digitization_of_analog_formats.pdf
http://rucore.libraries.rutgers.edu/collab/ref/dos_avwg_audio_obj_standard.pdf

Smithsonian Institution Archives
http://siarchives.si.edu/records/electronic_records/records_erecords_digitization_images.html

Stanford University
http://lib.stanford.edu/node/8544/moving-image-digitization
http://lib.stanford.edu/node/8544/audio-digitization

State Library of Queensland
http://www.slq.qld.gov.au/__data/assets/pdf_file/0009/139815/SLQ-_DS2Capture_v2.05.pdf

University of Illinois at Urbana-Champaign
http://images.library.uiuc.edu/resources/digitalguidev3.pdf

University of Maryland Libraries
http://www.lib.umd.edu/dcr/publications/best_practice.pdf

University of Michigan
http://www.lib.umich.edu/files/UMichDigitizationSpecifications20070501.pdf
http://www.lib.umich.edu/lit/dlps/dcs/UMichDigitizationSpecifications20070501.pdf

University of Southern California
http://www.usc.edu/libraries/collections/digitallibrary/documents/USCDL_Formats_and_Resolutions.pdf
http://www.usc.edu/libraries/collections/digitallibrary/documents/USCDL_Audiovisual_Digitization_Overview.pdf

University of Southern Mississippi Libraries Digital Program
http://www.lib.usm.edu/legacy/spcol/crda/guidelines/index.html

University of Virginia
http://etext.lib.virginia.edu/services/helpsheets/scan/specscan.html

University of Wisconsin
http://uwdcc.library.wisc.edu/documents/AudioWorkflow.pdf

Yale University
http://www.yale.edu/digitalcoffee/downloads/DigitalCoffee_SharedPractices_%5Bv1.0%5D.pdf

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

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).