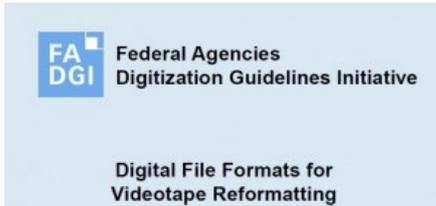# Comparing Formats for Video Digitization

Author: Carl Fleischhauer, a Digital Initiatives Project Manager in the Office of Strategic Initiatives

Source: http://blogs.loc.gov/digitalpreservation/2014/12/comparing-formats-for-video-digitization/

FADGI format comparison projects.] The Audio-Visual Working Group within the Federal Agencies Digitization Guidelines Initiative recently posted a comparison of a few selected digital file formats for consideration when reformatting videotapes. We sometimes call these target formats: they are the output format that you reformat to.

This video-oriented activity runs in parallel with an effort in the Still Image Working Group to compare target formats suitable for the digitization of historical and cultural materials that can be reproduced as still images, such as books and periodicals, maps and photographic prints and negatives. Meanwhile, there is a third activity pertaining to preservation strategies for born-digital video, as described in a blog that will run on this site tomorrow. The findings and reports from all three efforts are linked from this page.
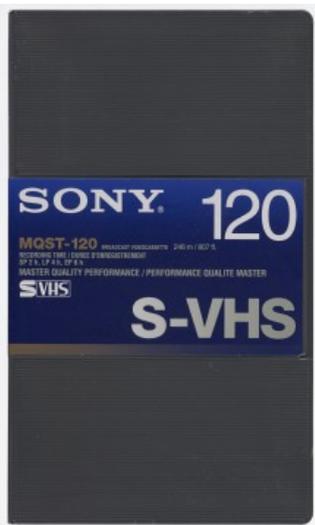
Courtney Egan, photo courtesy of NARA.

**Comparing video formats for reformatting.** The focus for this project was the reformatting of videotapes with preservation in mind, and it was led by Courtney Egan, an Audio-Video Preservation Specialist at the National Archives. Like its still-image parallel, the for-reformatting video comparison used matrix-based tables to compare forty-odd features that are relevant to preservation planning, grouped under the following general headings:

- Sustainability Factors
- Cost Factors
- System Implementation Factors (Full Lifecycle)
- Settings and Capabilities (Quality and Functionality Factors)

The online report offers separate comparisons of file wrappers and video-signal encodings. As explained in the report's narrative section, the term wrapper is ?often used by digital content specialists to name a file format that encapsulates its constituent bitstreams and includes metadata that describes the content within. A wrapper provides a way to store and, at a high level, structure the data; it usually provides a mechanism to store technical and descriptive information (metadata) about the bitstream as well.? The report compares the following wrappers: AVI, QuickTime (MOV), Matroska, MXF and the MPEG ad hoc wrapper.

In contrast, the report tells us, an encoding ?defines the way the picture and sound data is structured at the lowest level (i.e., will the data be RGB or YUV, what is the chroma subsampling?). The encoding also determines how much data will be captured: in abstract terms, what the sampling rate will be and how much information will be captured at each sample and in video-specific terms, what the frame rate will be and what will the bit depth be at each pixel or macropixel.? The report compares the following encodings: Uncompressed 4:2:2, JPEG 2000 lossless, ffv1, and MPEG-2 encoding.



S-VHS tape box, one of formats likely to be reformatted. Photo courtesy of NARA.

Courtney's team identified three main concepts that guided the analysis. First, the group sought formats that could be used to produce an authentic and complete copy of the original. An authentic and complete copy was understood to mean retaining specialized elements that may be present in the original videotape, e.g., multiple instances of timecode or multiple audio tracks, and metadata about the aspect ratio. Second, the group sought formats that maximized the quality of reproduction for both picture and sound. In general, this prejudiced the team against encodings that apply lossy compression to the signal.

Third, the group sought formats with features that support research and access. Central to this?especially for collections of broadcast materials?is the retention of closed captions or subtitles. These textual elements can be embedded in the file that results from the reformatting process and the text can later be extracted by an archive to, say, support word-based searching.

The desiderata of authentic copies and maximal support for research led Courtney's team to pay special attention to some fairly arcane technical factors. I'm not going to do much explaining in this blog (there's lots of good information online) but I will offer the following checklist to provide a sense of some techy elements that the team tracked as they made their comparisons:

- Bit Depth. This is a feature of encoding and, in the interest of quality, the team looked to see if higher-resolution 10-bit sampling was supported.
- Chroma Subsampling. For encodings, the team asked which forms of subsampling are supported? (Some provide higher quality than others.) For wrappers, the team asked, ?Is the type of subsampling in ?this file' declared in embedded metadata??
- Audio Channels. How many channels? Declared and tagged in metadata?
- Video Range. Does this format carry the ?rule-bound? broadcast range of luma and chroma data, or an unregulated ?wide range? signal that may have come from computer graphics? Is the range declared in embedded metadata?
- Timecode. Can multiple timecodes can be stored?
- Closed-captioning and Subtitles. Is there a specified location for captions in the file? Or must users employ sidecar files to retain this data?
- Scan Type and Field Order. Does this format support both interlaced-scan and progressive imagery? Is that fact (and also the field order for interlaced picture) declared in embedded metadata?

- Display Aspect Ratio. Is aspect ratio declared, specifically display and pixel aspect ratio?
- Multipart Essences. Support for segmentation, multipart essences?
- Fixity Checks. Does the format support for within-file fixity data? Many specialists wish to carry a fixity value for each video frame.



1-inch open reel tape. Photo courtesy of NARA.

Out of all of the comparisons, is there a single winning format? The team said no. Practices and technology for video reformatting are still emergent, and there are many schools of thought. Beyond the variation in practice, an archive's choice may also depend on the types of video they wish to reformat. The narrative section of the report indicates that certain classes of items?say, VHS tapes that contain video oral history footage?can be successfully reproduced in a number of the formats that were compared. In contrast, a tape of a finished television program that contains multiple timecodes, closed captioning, and four audio tracks will only be reproduced with full success in one or two of the formats being compared.

It is also the case that practical matters like an organization's current infrastructure, technical expertise and/or budget constraints will influence format selection. One of the descriptive examples in the narrative section notes, for example, that at one federal agency, the move to a better format awaits the acquisition of ?additional storage space and different hardware and software.?

**Sidebar: some preference statements suggest the existence of two communities.** The team talked to a variety of people as the work progressed and, in addition, sent copies of the final draft to experts for comment. As I reflected on the various contributions and comments the team received, I found myself pondering remarks about two lossless-compressed encodings: ffv1 and the ?reversible? variant of JPEG 2000. As far as I can tell, the two encodings work equally well: after you decode the compressed bitstream, you get back exactly what you started with, i.e., in both cases, the encoded data is ?mathematically lossless.? But each encoding had its own set of boosters. At great risk of oversimplification, I wondered if we were hearing from two different (albeit overlapping) communities, each with its own ethos.

One community is well represented by national libraries and archives, including the Library of Congress. When members of this community (I'm one of them!) select formats for video mastering and preservation, we are strongly drawn to ?capital-S? (official) standards. (When we select video ?access? formats for, say, dissemination on the Web, different factors come into play, more like those embraced by the open source advocates described below.)

We participate in or follow the work of standards developing organizations like the Society of Motion Picture and Television Engineers and the European Broadcasting Union. Our collections include significant holdings produced by broadcasters, content with complex added elements like captions and subtitles, multiple timecodes, and other elements. Although our standards-oriented community has moved vigorously toward file-based, digital approaches, its members are more likely to build production and archiving systems from the ?top down,? and employ commercial solutions. Now: how did this standard-oriented community vote on the lossless encodings? They favored lossless JPEG 2000, a standard from the International Standards Organization and the International Electrotechnical Commission.

3/4? U-matic tape case. Photo courtesy of NARA.

And the other community? These were specialists?several in Europe?who are strongly drawn to open source specifications and tools. My sense is that members of this group are eager to embrace formats and tools that ?just work,? and they are less firmly committed to capital-S standards. (I can imagine one of them saying, ?Let's just do it ? we have no time to wait for lengthy standard-development and approval processes.?) Many open source advocates are bona fide experts, skilled in coding and capable of developing systems ?from the bottom up.? Meanwhile, some of them work in or on behalf of archives where the collections do not feature extensive broadcast materials but rather consist of, say, oral history or ethnographic recordings, or other content made by faculty or students in a university setting, absent added elements like closed captions. In their communications with the FADGI team, several from this community favored the lossless ffv1 encoding. The published specification for ffv1 is authored by Michael Niedermayer and disseminated via [FFmpeg](). [Wikipedia]() describes FFmpeg as ?a free software project that produces libraries and programs for handling multimedia data.? Worth saying: the FFmpeg project commands considerable respect in video circles.

The simplified picture in this sidebar is, um, good fodder for a blog. But I'll be interested to hear if any readers also sense community-based preferences like the ones I sketched, which extend well beyond the matter of lossless encodings.

**Back to the FADGI comparison: no silver bullet.** Although no single format warranted an unqualified recommendation, our experience in comparing formats has been instructive, highlighting trends and selection factors, and winnowing the number of leading contenders down to a handful. We found that format preferences for the reformatting of video remain emergent, especially when compared to the better-established practices and preferences associated with still imaging and audio.