VAT Archival Best Practices: Additional Technical Info

Many thanks to Wikipedia for some details and text found here. Text from these sections is available as a modified work under the Creative Commons Attribution-ShareAlike License .

1. DIGITIZING AUDIO AND VIDEO

There are two components to digitizing a signal, discretization and quantization. In general, discretization and quantization occur at the same time, though they are conceptually distinct.

Discretization is the reading of an analog signal, and, at regular time intervals (at a particular frequency), sampling the value of that signal at that point in time. Each such reading is called a **sample** and may be considered to have infinite precision at this stage.

Quantization

Samples are then rounded to a fixed set of numbers (such as integers), in a process known as quantization. The quantity of the fixed set of numbers used to specify the value of the quantized sample are represented by bits. The larger the set of numbers, (more bits) the better the reproduction of the original signal.

Audio - An electrical signal analogous to a sound wave is created when a sound compression wave or mechanical vibration (generally in air, but solids and other fluids work as well) hits a microphone. Audio is a single waveform and the superposition of all the contributing audio sources add up to just one signal hitting the microphone. Most microphones today use electromagnetic induction (dynamic microphone), capacitance change (condenser microphone), piezoelectric generation, or light modulation to produce an electrical voltage signal from mechanical vibration. This electrical voltage signal is often still analog.

When digitizing, we talk about **raw audio** being sampled at a particular **frequency measured in KHz** (Kilo - thousands of Hz - cycles per second) with a particular resolution in **bits per**

channel of audio. 16 Bit 48 KHz Stereo Audio for example means that every 1/48,000 of a second the voltage or ”magnitude” or ”height” or “value” of the analog Audio waveform on both channels (left and right) is **sampled** and both are **quantized** to a number that is 16 bits long. The series of these measurements, when strung together in time comprise the new digital audio signal.

In 1 second, 48,000 measurements of 16 bits are taken,48,000 * 16 bits = 768000 bits * 1 Byte / 8 bits = 96,000 Bytes or raw audio 96 KB (KiloBytes) in size per channel of audio, 192KB in Stereo.

The **sample frequency** relates to **discretization**, and the **bits per channel** relate to the **quantization**. Together they determine the quantity of raw data that the signal is stored as digitally.

Video - The light wave passing through the lens of a digital camera and then interacting with a charge couple device (CCD), causes the collapse of that particular light wave's energy into a photon, which hits a particular physical section of the CCD with a particular amount of energy. When many many photons are registered this way on the CDD an image frame can be recorded as a huge number of electrical signals. CCD image sensors are widely used in professional, medical, and scientific applications where high-quality image data is required. In applications where a somewhat lower quality can be tolerated, such as webcams, cheaper active pixel sensors (CMOS) are generally used. When digitizing **raw images**, **quantization** is related to the bits per pixel (color depth) with which samples are recorded**,** the physical geometry of the CCD or CMOS sensor elements, and importantly, the number of those sensor elements (Megapixels). This can also be influenced by the pixel geometry when assembling the frame, and any hardware or software algorithms used to assemble the frame, as well as scanning type (progressive or interlaced). **Discretization** for video is related to the **frames per second** (eg. 24, 29.97, 30 , 60), or the rate at which you are capturing full or partial frame images from the CCD sensor.Video Sampling.

So for example 1 second of interlaced video at a 720 pixel by 480 pixel resolution, at 29.97 fps, with a bit depth of 24 bits per pixel,
29.97 fps * (720 * 480) pixels * 24 bits * 1 Byte/ 8 bits = 31072896 Bytes or raw video ~ 31MB in size.

For an hour of raw video data at this rate, 31MB *60 seconds/min * 60 min/hr = 111600MB/Hr. or 111.6 GB/Hr.. **This is a lot of data for SD! and we want to record it! That is why we have codecs. Codecs are meant to reduce the data rate and storage requirements of the recorded digital signals by removing data that does not significantly change our perception of the image.**

A **codec** is a hardware device or computer program capable of encoding or decoding a digital data stream or signal. A codec is a program, either *implemented* in hardware via specialized silicon (a specialized computer chip)*, or in software* by a set of *algorithms*, (running on a general purpose computer processor chip) which can read or write files with particular specifications. A **codec** (the *program that does the **co**ding and **dec**oding*) should not be confused with the data itself, which is a "coding" of data by the codec (program) to fit a particular compression *format* or *standard* (a way of storing (file) or transmitting (streaming) the data).

There are thousands of audio and video codecs, ranging in cost from free to hundreds of dollars or more. Many multimedia data streams contain both audio and video, and often some metadata that permit synchronization of audio and video. Each of these three streams (audio, video, metadata) may be handled by different programs, processes, or hardware; but for the multimedia data streams to be useful in stored (file) or transmitted (stream) form, they must be encapsulated together in a container format.

The metadata we describe in this context is metadata about the particular arrangement of the bits and bytes of that data within the video file, not the human associated information related to a particular program or "coding", like who produced it, when, where, and what it's title is.

A **container** or **wrapper format** is a metafile format whose specification describes how different data elements and metadata coexist in a computer file. In the case of multimedia, there are many container formats, you can compare some of them here.
Different codecs have different attributes and intended purposes. They compress audio and video with different algorithms / hardware. Some target lowest bit-rates or highest quality, others are meant for editing or recording, some use complicated predictive techniques to minimize data size.

2. STORAGE FORMATS

Popular video/audio codecs and wrappers for delivery and transmission:

- MPEG-4 good for online distribution of large video
- MPEG-2 used for DVDs, Super-VCDs, and many broadcast television formats
- MPEG-1 used for video CDs
- H.261
- H.263
- H.264 / AAC or mp3 **.mp4** lso known as *MPEG-4 Part 10*, or as *MPEG-4 AVC*, used for Blu-ray Discs and some broadcast television formats, patent encumbered. A new Cisco unencumbered encoder/decoder was released in 2013 - Mozilla is using it in Firefox for playback of .mp4 in HTML5 <video> <src> tags.
- H.265 / AAC also known as HEVC or the High Efficiency Video Codec, touted as providing twice the video compression as H.264 at the same PSNR, and slated as the successor to H.264 as we move into 4K UHD and 3D encoding scenarios.
- Theora / ogg-vorbis **.ogv** used for video on Wikipedia FOSS, and patent unencumbered
- Dirac / opus **.avi .mkv** encoder competing with H.264, Authored and maintained by the BBC.
- Daala / opus Next Generation open licensed video codec, successor to Theora, also known as NETVC, targeting UHD and 3D applications.

- VP8 / ogg-vorbis **.webm** made FOSS by Google when they purchased On2. It is likely patent unencumbered.
- VP9 / opus **.webm** Next generation codec by Google, successor to VP8, targeting 4K UHD and 3D applications.

Encoder comparisons:

http://keyj.emphy.de/video-encoder-comparison/

https://people.xiph.org/~xiphmont/demo/daala/update1.shtml

Some formats support chunked delivery of Video and Audio from segmented sets of video renditions for Adaptive Bitrate Streaming with a proper manifest file. here are some of the chunking standards:

- MPEG DASH - Industry Consortium
- HDS - Adobe
- HLS - Apple
- SIlverlight - Microsoft

Tapes and/or Camera Codecs (not exhaustive)

- Betacam SX, Betacam IMX, Digital Betacam, or DigiBeta — Commercial video systems by Sony, based on original Betamax technology
- AVCIntra 50 and 100 Panasonic format less space than DVCPROHD, comparable Quality.
- HDCAM was introduced by Sony as a high-definition alternative to DigiBeta.
- D1, D2, D3, D5, D9 (also known as Digital-S) — various SMPTE commercial digital video standards
- DV, MiniDV — used in older videotape-based consumer camcorders; designed for high quality and easy editing; can also record high-definition data (HDV) in MPEG-2 format

- DVCAM, DVCPRO — used in professional broadcast operations; similar to DV but generally considered more robust; though DV-compatible, these formats have better audio handling.
- DVCPRO50, DVCPROHD support higher bandwidths as compared to Panasonic's DVCPRO.
- Digital8 — DV-format data recorded on Hi8-compatible cassettes; largely a consumer format
- MicroMV — MPEG-2-format data recorded on a very small, matchbook-sized cassette; obsolete
- D-VHS — MPEG-2 format data recorded on a tape similar to S-VHS