Digitizing Sound Recordings and Moving Images Part Two: Let's Get Technical!

By Paul Eisloeffel, Nebraska State Historical Society

Editor's Note: The following article on digitizing audiovisual materials is the second in a series of three on the topic. Part one, "Why Digitize?" appeared in the January 2006 issue of the *MAC Newsletter*, and part three will follow in the next issue.

As you will recall, our last installment in this series explored the motivations for and benefits of digitizing sound recordings and moving images. We learned that these singular documents, in all their glorious variety, are frequent fixtures of our archival holdings and are increasingly in demand by our users. We also discovered that their dependence on machines to access their content fundamentally requires reformatting onto new media, and that digitization is increasingly the method of choice for reformatting analog audiovisuals and for copying those that are born digital.

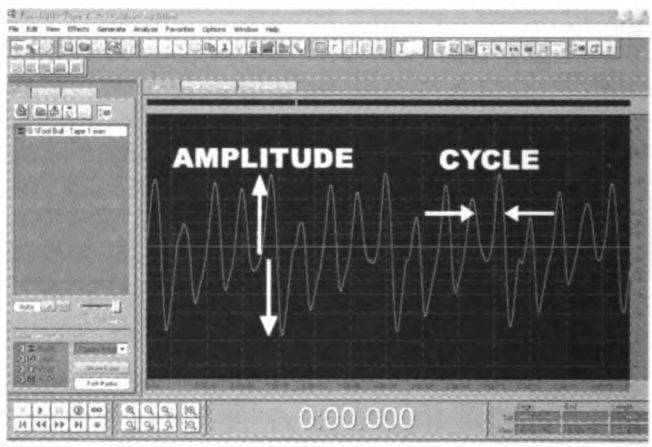
Many of us already have a lot of experience with flat scanning, and are comfortable with the terms and processes. But venturing down this *other* binary path may still fill some with trepidation. Indeed, digitizing audiovisuals is as different from flat scanning as the original documents are from one another. What do you need to know to understand the process?

To really get a handle on what happens when sounds and moving images are turned into a series of ones and zeros (digitized), you should first know something about the nature of their most elemental building blocks: sound and light.

The Anatomy of Natural Sound and Light

Sound and light both occur in nature in ways that are best represented by *waveforms*. In fact, the term *analog* refers to media formats and electronic devices that use waveforms to represent sound, light, and color, thus they produce a signal analogous to the natural phenomena.

Sound starts with a vibration that disturbs air molecules, creating differences in air pressure. The resulting movement of the molecules and variations in air pressure make up the waveform, the peak representing the high in the air pressure and the trough representing the low. *Amplitude* refers to the total change in pressure from the peak to the trough. Each pair of contiguous peaks and troughs makes up one *cycle*, the length of which is called the *wavelength*. *Frequency* is the number of cycles passing a specific point per second, expressed in kilohertz (kHz). The faster the wave is moving (vibrating), the more cycles pass by per second and therefore the higher the frequency.



This screen from the audio software program Audition shows a portion of the waveform of a digitized oral history recording, captured in mono at a sample rate of 44100 and bit rate of 16. This portion lasts only 26/1000 of a second. Notations pointing out amplitude and cycle have been added.

When a sound is recorded on analog equipment, the waves in air pressure are changed to waves of voltage, which make magnetic signals of various strengths on tape (or cut grooves of varying magnitude in disks, belts, or cylinders). Amplifiers and speakers turn these signals back into vibrations, and they go to our ears.

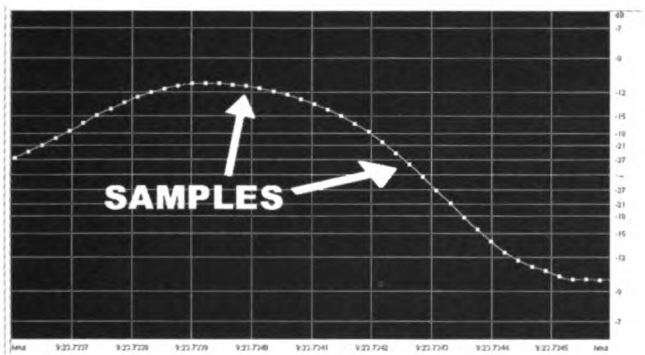
Light, the most obvious element of a moving image, is a lot more complicated, but its fundamental representation as a waveform is pretty similar. It, too, has amplitude and frequency, along with other elements such as luminance, color, areas of dark and light, and so forth. These all come together to represent a visual image. A *moving* image is a sequence of these stills, displayed one after another at sufficient speed that they are perceived by the human brain as blending into seamless motion. And in some cases, these images will be accompanied by corresponding sound.

Generated at Iowa State University on 2021-02-11 18:53 GMT / https://hdl.handle.net/2027/inu.30000107280145
Creative Commons Attribution-NonCommercial / http://www.hathitrust.org/access_use#cc-by-nc-4.0
Digitized by Google
Original from INDIANA UNIVERSITY

The Anatomy of Digital Audio and Moving Images

Keep these factoids about natural sound and light in mind, as they will help you understand what makes up digital audio and moving image files. First, digital audio.

One factor that determines the quality of a digital audio file is its *sampling rate*. The original analog waveform is digitally sampled at various points along its amplitude. The sampling rate refers to how often a sample is taken per second. For example, CD-quality sound is sampled at 44.1 kHz, or 44,100 times per second. The higher the sampling rate, the closer the digital waveform will resemble the original analog waveform. It's like a connect-the-dots drawing: The more dots there are, the smoother the line and the more detailed the picture. Put simply, the more samples, the better the sound.



Further detail of 1/1000 of a second shows the digital samples taken along the waveform.

Another factor that determines an audio file's quality is its *bit rate*, also known as *bit depth* or *sample size*. The bit rate refers to the amount of digital information devoted to describing what's happening at any given sample along the waveform—its *dynamic range*. The higher the bit rate, the more accurately the digital sampling reflects all the complexity and nuances of the original sound. To compare it to flat scanning, an audio file's bit rate is pretty much the same as the *resolution* of a still image file. CD-quality sound (used for music) is encoded at a rate of 16 bits. Besides the sampling rate and bit rate, digital audio files also have *channels*. These determine whether a recording is monophonic (one channel) or stereophonic (two channels).

A digital moving image file is a bit more complex. Its quality is additionally dependent on its *frame rate*, *data rate*, and *frame size*. The frame rate refers to the number of frames per second the moving image plays. The data rate, somewhat like an audio file's bit rate or a flat scan's resolution, indicates the amount of digital information devoted to describing the image at a given moment (like color depth, luminance, and contrast). Finally, because it is visual, a moving image file also has a frame size attached to it, just like a still image's spatial resolution or pixel dimension. Most frequently it is an aspect ratio of 4:3. By way of example, DVD-quality moving images play at about 30 frames per second (actually 29.97 fps), with a data rate of 5.7 megabits (Mb) per second and a frame size of 720x480 pixels per inch (ppi).

In addition to these factors, a digital moving image file may also include a sound component, adding to it all the elements of an audio file (sampling rate, bit rate, and number of channels).

Size *Does* Matter

In the most practical of terms, all the above-described complexity of digital audio and moving image files translates into their most significant feature: their enormous size. To put it into perspective, one minute of an audio file in the uncompressed WAV format containing stereo sound encoded at a sampling rate of 44.1 kHz and a bit rate of 16 (CD-quality) takes up 10 megabytes (MB) of space; one hour uses 600 MB. Moving images are larger still: one minute of video in DV format with a frame rate of 30 fps, a data rate of 5.7 Mb, and a frame size of 720x480 ppi fills more than 200 MB of space, and one hour commands more than 20 gigabytes (GB).

Even though advances in digital technology are ongoing, media files of such bulk still strain the limits of practical storage and use. As with flat scans, our goal should be to create the highest-quality digital surrogate possible, so as to allow the "retirement" of the original document. Nevertheless, storage of files this big can be a burden, and their use fairly well impossible. Consequently, decreasing their size is generally a necessary part of working with digital audio and moving image files.

Let's Get Small!

There are basically two ways to bring media files down to a practical size. The first is to reduce any of the above-mentioned factors that determine that media file's quality, not unlike reducing the resolution or pixel dimension of a flat scan. Cutting the sampling rate, bit rate, or number of channels of an audio file will reduce its size significantly; so too will trimming down the frame rate, data rate, or frame

(Continued on page 24)

Generated at Iowa State University on 2021-02-11 18:53 GMT / https://hdl.handle.net/2027/inu.30000107280145
Creative Commons Attribution-NonCommercial / http://www.hathitrust.org/access_use#cc-by-nc-4.0
Digitized by Google
Original from INDIANA UNIVERSITY

(Continued from page 23)

size of a moving image file. However, these reductions will have a decided effect on the quality of the files.

The more acceptable tool for bringing these huge files down to size is called *compression*. Like compression in other automation scenarios, media file compression is either *lossless* or *lossy*. When a file is compressed in a lossless fashion, all of the data remains but is squeezed into a smaller space for storage and delivery. When accessed, the file returns from compression unaltered.

Lossy compression, on the other hand, identifies and discards redundant data and irretrievably alters the file. It does this by playing off the limits of human perception, or the fact that we cannot discern auditory and visual stimuli beyond certain parameters, and replacing or eliminating the data that humans can't detect. In audio files this is called *psychoacoustic compression*, which removes sound above and below the range of human hearing or sounds that are masked by more dominant ones. Lossy moving image compression utilizes *psychovisual* techniques, which remove redundancies based on the human eye's inability to discern fine spatial detail, detail near object edges, or slight changes from frame to frame.

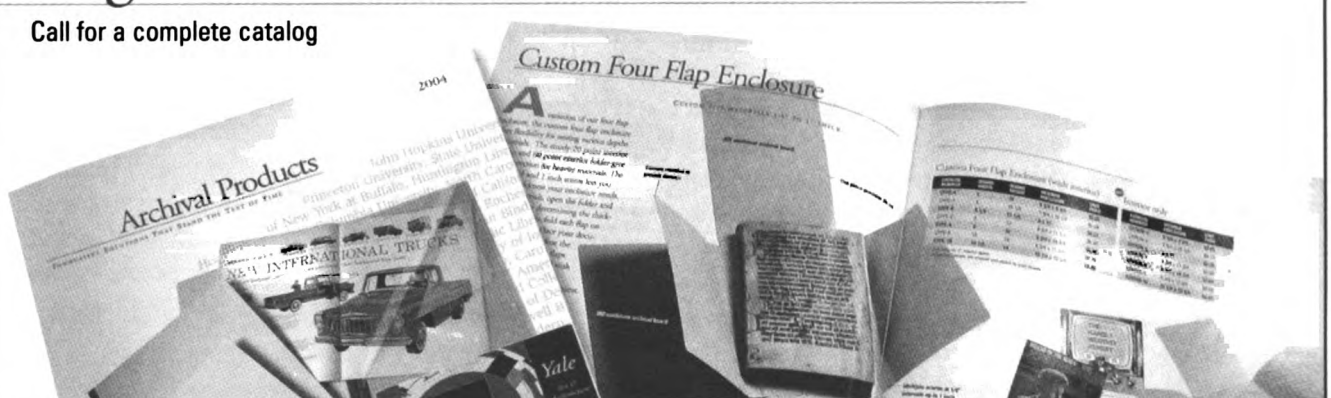
The mechanism through which media file compression is accomplished is called a *codec*. A codec (*COmpressor/DECompressor*) is a mathematical algorithm through which a file is filtered. This algorithm *compresses* the file for storage and delivery, and then *decompresses* it when it is accessed. There are many codecs, some designed expressly for audio files and some for moving images. Most are particularly suited for a specific application, like putting music files on CDs, burning video onto DVDs, videoconferencing, and talking on a digital cellular phone. Current media codecs can achieve more than a 50:1 reduction in file size.

As you can see, the world of digital sound recordings and moving images can be a complicated place! For better or for worse, much of this complexity is hidden behind the ever-increasing user-friendliness of multimedia hardware and software, and this may give us the impression that we don't truly *need* to understand what's really going on. However, as with all other things, we are always better practitioners of our craft when we are familiar with the fundamentals.

In our last installment of this series, we will explore the steps involved in creating and delivering a digital audiovisual product.

Elegant Solutions for Preservation

Call for a complete catalog



Pamphlet Binders
Music Binders
Archival Folders
Manuscript Folders
Hinge Board Covers
Academy Folders
Newspaper/Map Folders
Polypropylene Sheet & Photo Protectors
Conservation & Preservation Tapes
Bound Four Flap Enclosures
Archival Binders
Archival Boards
Adhesives
Bookkeeper
Century Boxes
Record Album Enclosures
Conservation Cloths

ARCHIVAL PRODUCTS
P.O. Box 1413 | Des Moines, Iowa 50305-1413
800.526.5640 | Fax 888.220.2397
custserv@archival.com | www.archival.com

archival.com

Generated at Iowa State University on 2021-02-11 18:53 GMT / https://hdl.handle.net/2027/inu.30000107280145
Creative Commons Attribution-NonCommercial / http://www.hathitrust.org/access_use#cc-by-nc-4.0

Digitized by Google

Original from INDIANA UNIVERSITY