

Marie Lascu
3403 Lacinak/Oleksik
12/14/2011

Comparing Audio Compression Rates

The goal was to assemble a diverse enough selection of audio samples from the fine collection of test materials in the MIAP lab room, and then create multiple digital files at different audio compression rates for comparison purposes. Based on materials available in the lab, two 1/4" open reel tapes and three audio cassette tapes were chosen.

Equipment

The 1/4" open reel audio was digitized using a Studer 807 playback machine which was plugged into a M Audio Firewire solo, a portable audio interface that provides 24-bit audio quality and is compatible with most music software. The Firewire facilitated digital conversion from the analog machine onto a computer desktop using Wavelab, a digital audio editing program. The audio cassette material was digitized using a Tascam 112 Mk II playback machine, which was also connected to the M Audio Firewire solo for conversion through Wavelab.

The listening tests were performed by connecting my Asus laptop to a Harman Kardon stereo receiver, and using Sennheiser HD555 head phones. The hope being that the stereo receiver would boost the sound through the headphones.

Methodology

Not being equipped with state-of-the-art studio equipment, I had to settle for listening to each sample through the equipment listed above. Each sample was listened to on the same exact equipment for consistency of sound quality.

I listened to the uncompressed file version of each sample 1-2 times to get a sense of the acoustic characteristics. I then listened to the compressed files in descending order (320 kbps

→224 kbps → 128 kbps → 96 kbps → 56 kbps) and took notes on what I heard as I was listening. The compression rates were chosen randomly. I used Audacity, another digital audio editing program, to play the files. It was very useful in terms of the visual waveform it provides which allowed me to pinpoint auditory nuances more efficiently for comparison purposes.

Results

A handy spreadsheet, too long to fit.

	Type	Format	File Name	Sample Path
Source 1	Spoken Word	1/4" open reel	spoken word open reel_24 96.wav	Studer 807 - M audio firewire solo - wavelab
Source 2	Music - Opera	1/4" open reel	music open reel_24 96.wav	Studer 807 - M audio firewire solo - wavelab
Source 3	Music - Mingus	audio cassette	music cassette_24 96.wav	Tascam 112 mk II - M audio firewire solo - wavelab
Source 4	Spoken Word	audio cassette	spoken word cassette_24 96.wav	Tascam 112 mk II - M audio firewire solo - wavelab
Source 5	Music - Waits	audio cassette	music 2 cassette_24 96.wav	Tascam 112 mk II - M audio firewire solo - wavelab

Cont'd

Uncompressed	320 kbps	224 kbps	128 kbps	96 kbps	56 kbps
can hear very mild fuzz noise in background, S pronounced but not terribly prominent	similar	similar	similar	background fuzz has more metallic sound, S sound is more pronounced	more distant fishbowl sound, more metallic noise
mild fuzz noise in background, otherwise clear vocals, small pop just past 1:15 (in room or transient?), background fuzz increases near end of clip	similar	similar	similar	hearing more metallic frequency in background- less fuzz	immediately sounds more distant, like in a fishbowl, fuzz has become total computer sounding frequency, horns very muffled, pronounced S

alot of instrumentation- crisp sounding	similar	similar	similar	horns sounded slightly less loud, very slight	slight fishbowl sound, more distant
poor quality, alot of interference, voice sounds extremely distant, music randomly comes in at about 35 sec	similar	similar	interference a bit more prominent, louder	getting more difficult to hear speaking voice	the quality is so bad to begin with, sounds a little more awful
some fuzz frequency audible, but not too prevalent	similar	similar	similar	possible slight decrease in volume, not sure, fuzz in background slightly more prevalent, very slight	big change, fishbowl faraway sound, general mechanical noise in background

○ Source 1: Spoken Word (radio interview), 1/4" open reel

The uncompressed 24bit 96 Hz .wav file contained very mild fuzz in the background. The interviewee had slightly pronounced Ss, but it was not an extreme example.

- 320 kbps
I could not audibly detect a discernable difference in sound.
- 224 kbps
I could not audibly detect a discernable difference in sound.
- 128kbps
I could not audibly detect a discernable difference in sound.
- 96kbps
The earlier detected background fuzz has taken on a more mechanical sound, and the S sound in the interviewee is also more pronounced with a mechanical tinge.
- 56 kbps
There is an immediately noticeable disparity in quality, the audio has a more distant fishbowl sound, and the background noise sounds even more mechanical.

○ Source 2: Music (opera), 1/4" open reel

The azimuth was adjusted before recording and was Chris approved. The uncompressed 24bit 96Hz .wav file also contained mild fuzz in the background, but otherwise the vocals were clear and so was the orchestral accompaniment.

- 320 kbps
I noticed a slight pop after the 1:15 mark, but it is unclear to me if it is a transient noise or an occurrence in the room the material was recorded in. Otherwise I could not audibly detect a discernable difference in sound.
- 224 kbps
I could not audibly detect a discernable difference in sound. The vocals continue to sound crisp.
- 128kbps

I could not audibly detect a discernable difference in sound.

- 96kbps
The earlier detected background fuzz took on a more mechanical sound, as with the previous sample.
- 56 kbps
There is an immediately noticeable disparity in quality, the background fuzz sounds like a computer frequency, the horns in particular sound muffled. The vocals now have over-pronounced Ss.

○ Source 3: Music (Charles Mingus), audio cassette

The uncompressed 24bit 96Hz .wav file highlights the crisp sounding instrumentation and gruff vocals.

- 320 kbps
I could not audibly detect a discernable difference in sound.
- 224 kbps
I could not audibly detect a discernable difference in sound.
- 128kbps
I could not audibly detect a discernable difference in sound.
- 96kbps
The only difference I could discern was that the horns sounded a little less loud.
- 56 kbps
While not as great a drop as the first two samples, the music does sound slightly more distant.

○ Source 4: Spoken Word (interview), audio cassette

The uncompressed 24bit 96Hz .wav file comes from a source that was already of poor muffled quality. There is a great deal of interference, and the lone voice on the recording is extremely distant. Music also randomly pops in at the 35 sec mark for a mere 2 sec.

- 320 kbps
I could not audibly detect a discernable difference in sound.
- 224 kbps
I could not audibly detect a discernable difference in sound.
- 128kbps
The interference sounds more prominent.
- 96kbps
At this point it is becoming difficult to hear the already compromised speaking voice.
- 56 kbps
Presumably worse, but the quality is so awful to begin with.

○ Source 5: Music (Tom Waits), audio cassette

The uncompressed 24bit 96Hz .wav file contains some audible fuzz, but it is not too prevalent. Otherwise it sounds like normal raspy Tom Waits.

- 320 kbps
I could not audibly detect a discernable difference in sound.
- 224 kbps
I could not audibly detect a discernable difference in sound.
- 128kbps
The interference sounds louder, more prominent.
- 96kbps
There may have been a slight decrease in volume, but I'm uncertain. The background fuzz is slightly more prominent.
- 56 kbps
There is an immediately noticeable disparity in quality, that fishbowl distant sound, and the fuzz in the background has a more mechanical tone.

Conclusion

The definition of compression alone (The reduction of data in a recorded waveform for the purpose of transmission) makes it an ant-archiving concept. What was interesting is that consistently the first three compression rates I listened to (320 kbps, 224 kbps, and 128 kbps) had little difference in an audible sense, but the very fact that each is transmitting less data than the one prior is disturbing.

For the purposes of preservation, there is no question about which compression rates are acceptable because the answer is none. I could not tell the difference between most 128 kbps files in comparison with the original, but if that 128 kbps copy were to outlive the source (which will undoubtedly occur) than all subsequent copies would have to be made from this inferior surviving copy, thus creating a slew of even lesser quality copies, making for a very unfortunate archive. In terms of access copies, however, I fully support using the highest rate of compression that does not compromise the user's listening experience. This simply makes sense for conservation of digital storage space, and new access copies can be made any time from the higher quality uncompressed digital audio file.