Lossless audio compression standards

Andrei Buruntia Paul Cazan

21/03/2018

Outline

- Terminology
  - Compression
  - Compressor controls
- Basic principles of lossless compression
  - Framing
  - Intra-channel decorrelation
  - Entropy coding
- Performance evaluation

Outline

Terminology

Compression

- Compression is the process of lessening the dynamic range between the loudest and quietest parts of an audio signal. This is done by boosting the quieter signals and attenuating the louder signals.

Outline

Terminology
- Compression
- Compressor controls

Basic principles of lossless compression
- Framing
- Intra-channel decorrelation
- Entropy coding

Performance evaluation

Terminology

Compressor controls

- Threshold - how loud the signal has to be before compression is applied.
- Ratio - how much compression is applied. For example, if the compression ratio is set for 6:1, the input signal will have to cross the threshold by 6 dB for the output level to increase by 1dB.
- Attack - how quickly the compressor starts to work.
- Release - how soon after the signal dips below the threshold the compressor stops.

Terminology

Compressor controls

- Knee - sets how the compressor reacts to signals once the threshold is passed. Hard Knee settings mean it clamps the signal straight away, and Soft Knee means the compression kicks in more gently as the signal goes further past the threshold.
- Make-Up Gain - allows you to boost the compressed signal. as compression often attenuates the signal significantly.
- Output - allows you to boost or attenuate the level of the signal output from the compressor.

Compression

- The figure is a block diagram representation of the operations involved in compressing a single channel.
- Generally, stereo or multiple channels are compressed separately.

Figure 1 The basic operations in most lossless compression algorithms.

Outline

Basic principles of lossless compression

Framing

- Framing provides editibility, a very important property for the signale to have in applications dealing with digital audio.
- It is important to quickly be able to edit a compressed bit stream.
- For practical purposes, the signal is divided into time frames of equal length.
- Each frame has a header containing the compression parameters. These can change on a framte-to-frame basis to follow the changing characteristics of the signal.
- If the frames are too short, the headers can generate a significant overhead.
- If the frames are too long, we lose editibilty.
- The most commonly used algorithms use a 13-26 ms time frame durations, translating into 576-1152 samples at 44.1 kHz.

Outline

The basic principles of lossless compression

Intra-channel decorrelation

- The purpose of this step is to remove redundancy by decorrelating the samples within a frame.
- 2 approaches:
  1. Modified linear predictive modelling of the signal.
  2. Obtaining a low bit-rate quantized or lossy representation of the signal, copmute the difference between the signals and losslessly compress the copmuted signal.

Outline

The basic principles of lossless compression

Entropy coding

- Entropy coding removes redundancy from the residual signal.
- No information is lost during the process.
- Three widely used methos: Huffman coding, run length coding, Rice conding.
- We will discuss the less known of the three, Rice coding.

Entropy coding

Rice coding

- Rice coding is characterized by one parameter (m).
- In this type of coding, a number is divided into 3 parts: a sign bit, the m low-order bits, and the remaining higher-order bits.
- The bit sign is self-explanatory, the second part consists of the m least significant beats of the binary represantation of the number's absolute value, and the third part is made of N consecutive zeros, where N has the same binary representation of the yet unused most significant bits; a 1 bit is inserted to terminate the sequence.
- The table below show the Rice code representation for parameter m = 3.

Number	Sign bit	m Lower bits	Number of 0's	Full code
0	0	000	0	00001
18	0	010	2	0010001
-12	1	100	1	110001

Performance evaluation

Test suite

- All of the tracks are in CD format (44.1 kHz, stereo, 16 bits).

- Track 4 of [17] (3 min 11 s): Madonna's "Like a Virgin" song.
- Track 20 of [7] (39 s): Saxophone, low frequency musical signal.
- Track 27 of [7] (20 s): Castanets, high frequency musical signal.
- Track 48 of [7] (28 s): Voice quartet.
- Track 66 of [7] (18 s): Wind ensemble.
- Track L=R (18 s): this track was constructed using the left channel of Track 66 of [7]. The right channel is an exact copy of the left channel. Therefore, the difference between the left and right channel is zero. This track serves as a probe to allow us to determine if a given compression scheme takes advantage of inter-channel correlation.

Performance evaluation

Results

Track	File size (bytes)	H_0 Left channel (bits per sample)	H_0 Right channel (bits per sample)
Track 04	33,715,920	14.33	14.28
Track 20	6,879,600	10.86	11.03
Track 27	3,528,000	9.27	9.26
Track 48	4,939,200	11.59	11.48
Track 66	3,175,200	10.81	10.92
Track L=R	3,175,200	10.83	10.83

Table 3 File size and first-order entropy for left and right channels ($x[n]$).

Performance evaluation

Results

Track	Integer arithmetic	Fixed-point arithmetic	
	AudioPaK	MUSICompress	Sonarc
Track 04	1.39	1.37	1.42
Track 20	3.12	2.93	3.13
Track 27	2.47	2.46	2.71
Track 48	2.56	2.41	2.65
Track 66	2.46	2.33	2.55
Track L—R	4.93	4.58	2.56

Table 4 Compression ratios for integer and fixed-point arithmetic lossless audio coders (we set the frame size to 1152 samples for the integer coders, and we used the default arguments for the fixed-point coders).

Performance evaluation

Results

Track	Floating-point arithmetic					
	Shorten -p0	Shorten -p10	OggSquish	LTAC	Sonarc -x	WA -c5
Track 04	1.38	1.43	1.43	1.41	1.42	1.46
Track 20	3.11	2.72	3.01	3.11	3.16	3.28
Track 27	2.46	2.69	2.67	2.70	2.72	2.83
Track 48	2.54	2.32	2.53	2.65	2.69	2.77
Track 66	2.46	2.42	2.51	2.54	2.57	2.64
Track L=R	2.47	2.44	5.01	2.70	2.58	5.28

Table 5 Compression ratios for floating-point arithmetic lossless audio coders (we set the frame size to 1152 samples for Shorten and we used the default arguments for OggSquish, Sonarc, LTAC, and WA).

Performance evaluation

The data

- The data in tables 4 and 5 is very revealing. Firstly, we note that the compression ratios obtained by various state-of-the-art algorithms are quite similar.
- The data for Track R=L shows which of the algorithms attempts to take advantage of intra-channel dependencies in a way that works best when the two channels are tightly correlated (in this case they are the same).
- Finally, the tests show that WA with the -c5 consistently gives the best compression throughout the 6 tracks.

Summary

- This study found that the currently available compression technology has reached a limit in what we can achieve in lossless audio compression.
- Furthermore, in spite of the wide range of approaches and complexities of the differenet compressors, they all perform similarly on the same audio signals, generally with less than one bit per sample in difference.
- It is noteworthy that HP's AudioPaK has similar performance to other modern compressors while keeping the complexity and number of arithmetic operations down. It is likely that upcoming compressor designs will focus on reducing algorithm complexity.