

# Ogg Vorbis Audio Compression

## How It Works

Martin McCrory  
8 December 2005

Step 1: Original audio transformed using a Modified Discrete Cosine Transform (MDCT)

- waveform in the time domain is converted to a waveform in the frequency domain
- like a FFT, but the MDCT has the unique property of being “lapped” - e.g. All “snapshots” of sound overlap with each other, meaning that no sound is lost between each “snapshot”
- Also, a MDCT has half as many outputs as inputs, meaning the resulting set of snapshots takes up half as much space
- Additionally, Vorbis adds a “window function” to the MDCT to eliminate potential errors at the beginning and end of each window (if the sample doesn’t loop perfectly, the window function smooths the ends of the window, increasing the accuracy of the MDCT while not changing the size of the “snapshot”)

Step 2: Separate the signal into two components: a “noise floor” and the sounds we are actually interested in

- Noise floor: low-volume sounds we’re not too interested in encoding properly
- Everything else: sounds of higher volume that exceed the noise floor (these are the sounds that must be compressed accurately)

Step 3: Quantization and compression of sound

- Sound is already quantized, in a sense
- Aim of Vorbis compression is to come up with a way to re-quantize the sound such that the compressed file takes up less space
- **Solution: Use a Codebook!**
  - Codebook: database of symbols that represent all of the various sounds
  - Groups of symbols are used when no symbol exists for a given sound (much like the hierarchy of letters, words and sentences to form linguistic meaning)
  - Entropy coding: Codes are assigned according to the negative probability of their occurrence. Meaning, the most common sound have the most efficient codes.

Other pertinent notes:

- This codebook process results in a file that is not compressed at a constant

“bitrate,” but is compressed efficiently for the sounds in the sound file. However, bitrates can be approximated by adjusting the size of the “snapshots” of the MDCT and other means.

- Vorbis files compressed at too low a bitrate have noise floor leakage: the windows aren’t small enough to capture all of the instances of “important” sounds, and some noise floor sound is incorporated into the compressed file (analogous to the metallic “warbling” of overcompressed MP3 files)

Step 4: Decode the Codebook and put the pieces of sound back together!

- The Vorbis decoder has a copy of the Codebook, and simply puts the pieces of compressed audio back together according to the sounds represented by the codes!
- Not all data is encoded, and then decoded. Sounds from the noise floor, for example, and sounds outside of the human hearing range.