How Audio Compression Works — Cheat Sheet

The four ideas behind every codec, and which idea each codec leans on.

The four ideas

1 · Psychoacoustic masking	Deletes sounds hidden behind louder ones — the ear can't hear them. *Heavy lifter for MUSIC. Lossy.*
2 · MDCT (frequency transform)	Turns a slice of waveform into frequencies, with no storage penalty. *The tool that makes masking possible. Lossy setup.*
3 · Linear prediction (LPC/CELP)	Models the voice, stores only the small prediction error. *Heavy lifter for SPEECH. Lossy.*
4 · Entropy coding	Short codes for common values, long for rare. Throws nothing away. *Shared final stage. LOSSLESS — in every codec.*

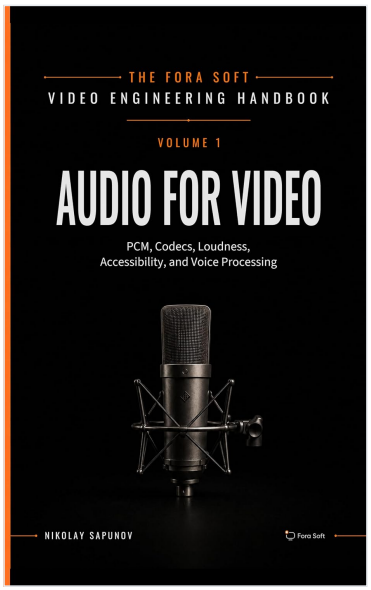
Which idea does each codec use?

Codec	Masking	MDCT	Lin. pred.	Entropy	Tribe
MP3 / MP2	Yes	Yes	No	Huffman	**Music**
AAC-LC	Yes	Yes	No	Huffman	**Music**
AC-3 (Dolby Dig.)	Yes	Yes	No	Custom	**Music**
Vorbis	Yes	Yes	No	Codebook	**Music**
G.711	No	No	No	None	**Speech**
CELP speech	Light	No	Yes	Varies	**Speech**
Opus — SILK mode	Light	No	Yes	Range	**Speech**
Opus — CELT mode	Yes	Yes	No	Range	**Music**
xHE-AAC (USAC)	Yes	Yes	Yes	Arithmetic	**Both**
FLAC / ALAC	No	No	Yes	Yes	**Lossless**

Read it this way

Transform + masking = a MUSIC codec. Linear prediction = a SPEECH codec. Both at once = a unifier (Opus, xHE-AAC). Entropy coding is always the last step. Lossy throws away the inaudible; lossless (FLAC) keeps everything and tops out near 2:1.

Companion to: 'How Audio Compression Works: The Four Ideas Behind Every Modern Codec'
www.forasoft.com/learn/audio-for-video/articles-audio/how-audio-compression-works



GO DEEPER · VOLUME 4 OF 7

Audio for Video: A Complete Guide to Sound in Video Products

From PCM and Codecs to Loudness, Accessibility, and Real-Time Voice Processing

By Nikolay Sapunov, CEO at Fora Soft · 1,229 pages · Kindle & paperback on Amazon

Read it on Amazon → amazon.com/dp/B0HH6PRZH1

This sheet reduces audio compression to four ideas, masking, MDCT, linear prediction and entropy coding, and shows which codec leans on which. Volume 4 develops each one in depth, from PCM, sampling and dynamic range to the Opus, AAC and speech codecs built on them.

WHAT YOU WILL FIND INSIDE

- Sampling, bit depth and dynamic range
- Speech coding with linear prediction
- Transform codecs: MDCT and masking
- Opus and AAC internals

THE FULL HANDBOOK · 7 VOLUMES · 9,000+ PAGES · ALL ON AMAZON

Vol. 1 · Video Encoding — pixels and color to H.264, HEVC, AV1, quality metrics and AI
Vol. 2 · Video Streaming I — network transport, contribution and ingest, HLS, MPEG-DASH, CMAF
Vol. 3 · Video Streaming II — ABR, CDNs, players, WebRTC, DRM and the economics of delivery
Vol. 4 · Audio for Video I — PCM, codecs, loudness, accessibility, real-time voice processing
Vol. 5 · Audio for Video II — jitter buffers, lip sync, Atmos, quality metrics, neural codecs
Vol. 6 · AI for Video I — models and inference, computer vision, speech, search, generative video
Vol. 7 · AI for Video II — real-time pipelines, agents, serving, cost, compliance, verticals

Tuning audio quality against bitrate in your product?

We will profile your content, set codec and bitrate per use case and verify the result with listening tests. 250+ video products shipped since 2005.

Book a 30-min call → **WhatsApp →** **Email us →**