Digital Audio Compression: Why, What, and How

An Absurdly Short Course

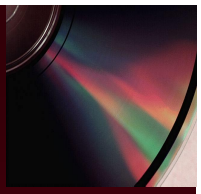
Jeff Bier

Berkeley Design Technology, Inc.

Outline

- Why Compress?
- What is Audio Compression?
- How Does it Work?
- Conclusions

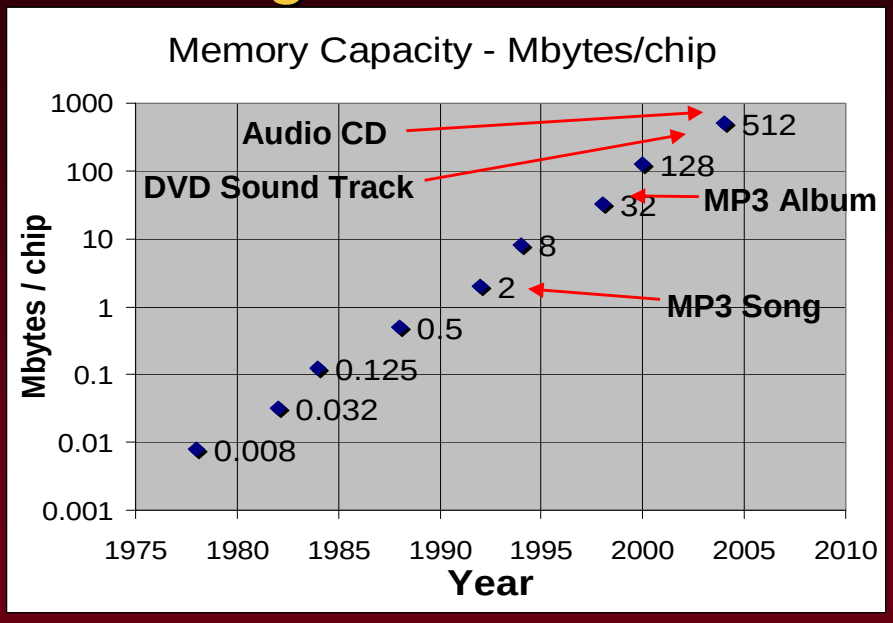
Why: Too Much Data!



- For example, Compact Disc
  - Rate: 2 channels * 44.1 ksamples/sec * 16 bits = 1.4 Mbit/second audio data
  - Audio CD capacity: 6.3 Gbits audio data
- 6-channel: 6 channels * 48 ksamples / sec * 16 bits = 4.6 Mbit/second
- Typical compressed audio today: a few hundred kbit/sec for even 6 channels

3

Crossing Thresholds



4

Commercial Applications

- Cinema (digital film sound)
- Consumer devices
  - Mini-disc (Sony)
  - Handheld players
  - Games
  - DVD
- Distribution
  - Internet distribution
  - Audio over cable (set-top box)
  - Satellite/terrestrial audio broadcast
  - Digital television

5

What: Compression Goals

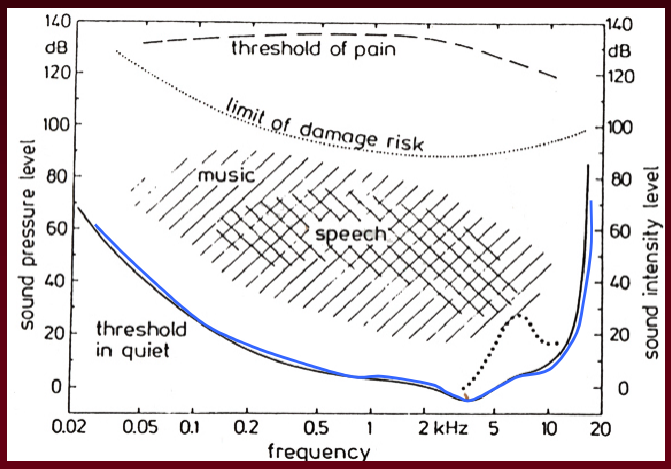
- Reduced bandwidth and/or storage
- Make decoded signal sound as close as possible to original signal
- Lowest implementation complexity
- Reasonable data-type requirements
- Applicable to as many signal types as possible
- Robust
- Scalable
- Extensible

6

How: Psychoacoustics

- What is it?
  - Relationship between what arrives at the ear and what we hear
- Why is it important for compression?
  - Don't transmit what the ear can't hear
  - Allow added noise that can't be heard
- What will we cover today?
  - Range of human hearing
  - Masking
    - Frequency
    - Time

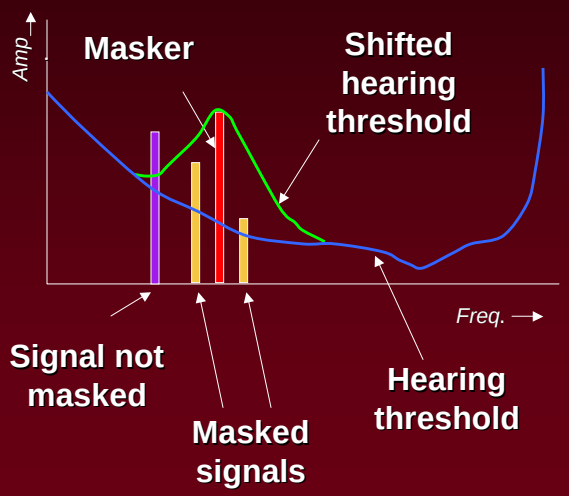
Range of Human Hearing



Zwicker/Fastl p. 17

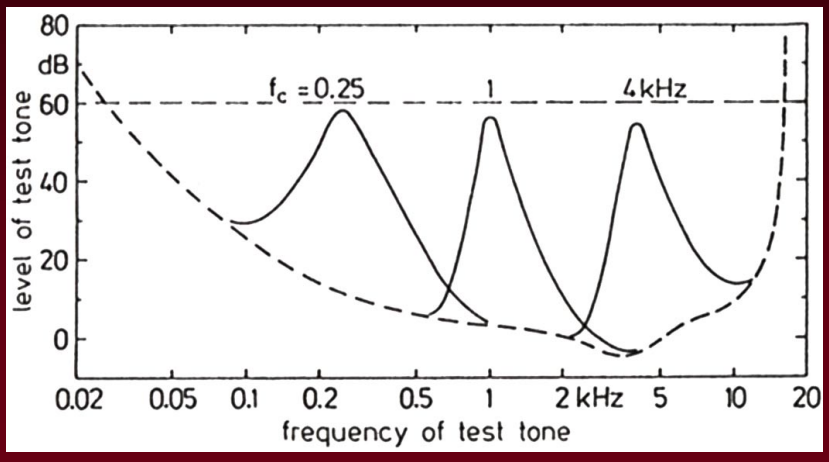
Auditory Masking

- One signal can make another inaudible

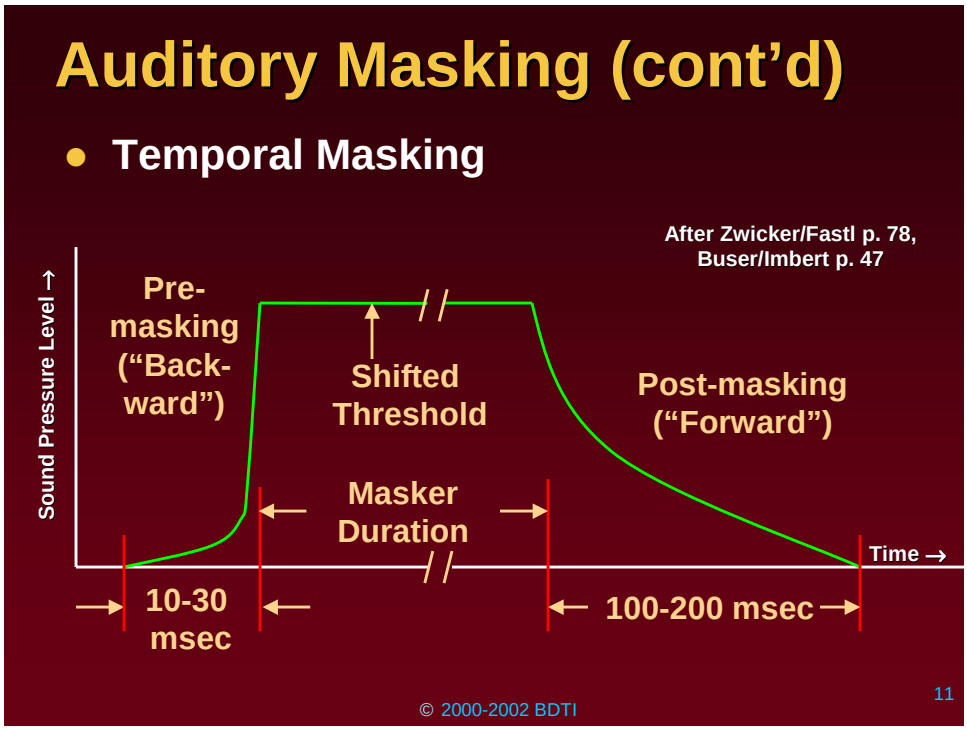


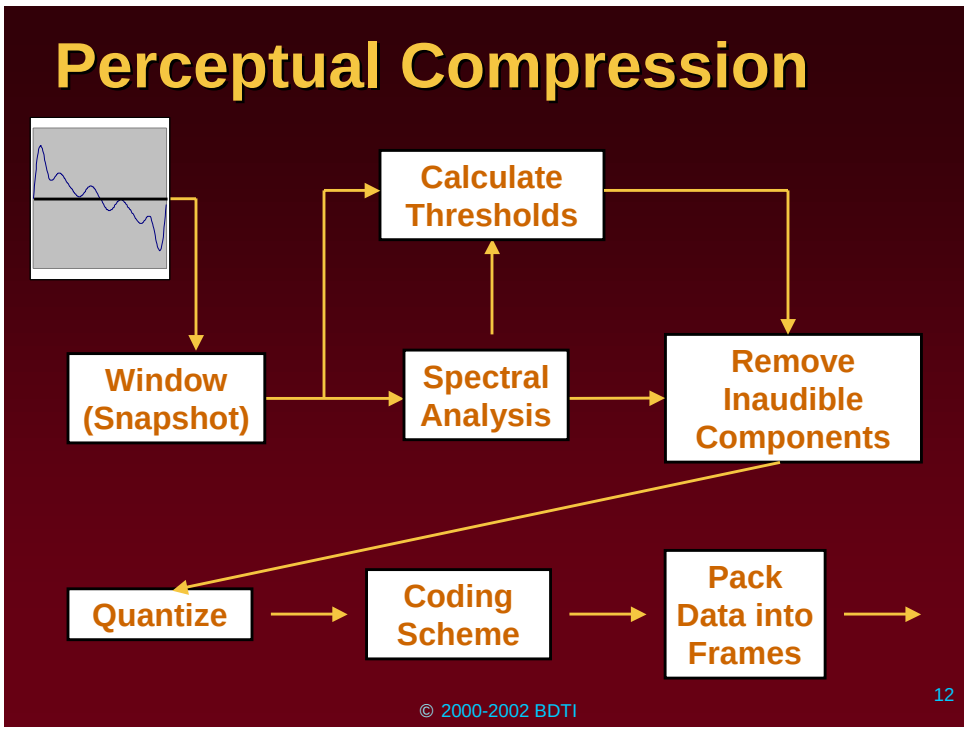
Auditory Masking (cont'd)

- Simultaneous masking, narrow-band noise masker



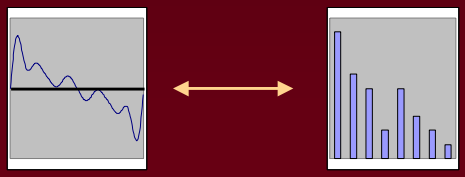
Zwicker/Fastl, p. 64





How: Spectral Analysis

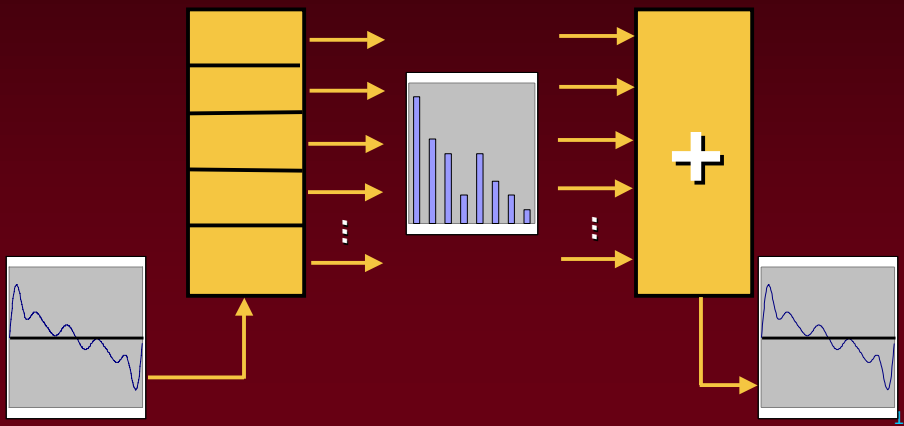
- What is it?
  - Break a signal into spectrum
  - Recover the signal from its spectrum



13

Spectral Analysis (cont'd)

- How is it done?
  - Filter bank
  - DCT



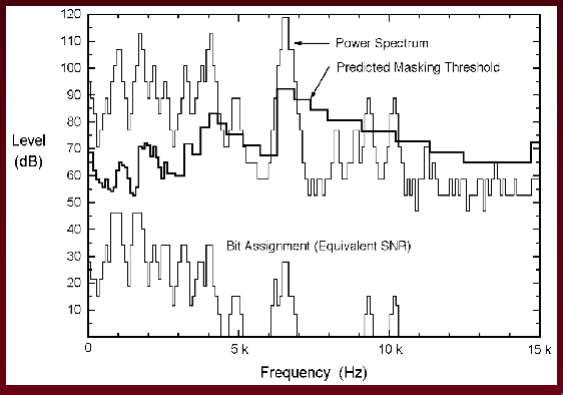
14

Spectral Analysis (cont'd)

- Analysis/resynthesis can be an "identity system" with proper constraints:
  - Design of window
  - How often take transform
  - How to overlap windows in analysis/resynthesis

How: Noise Allocation

- Quantize coefficients for audible spectral components to minimum # bits
- Adds noise
  - Keep below masking threshold



Davidson et al., 1994

How: More Tricks

- Filter
  - Band-limit input signal
  - LFE bandlimited to <120 Hz
- Differential coding of spectral values
- Coupling
  - Across time
  - Across channels ...

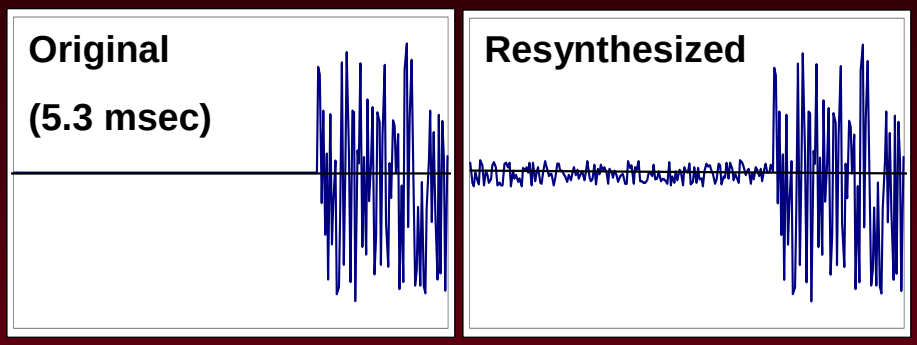
17

How: Data Compression

- Entropy (e.g. Huffman) coding
- Quantization table
- Run-length
- Vector quantization

18

Artifacts: "Pre-echo"



- Quantization noise spread
- Noise components ≥ 1-2 msec before impulsive signal not masked
- Fix: Shorter window; wavelets (ePAC)

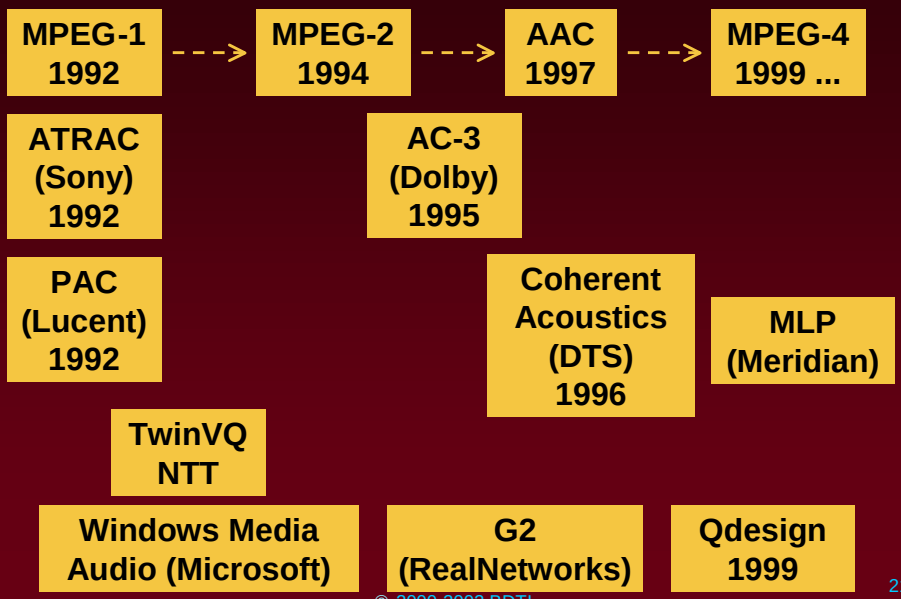
19

Comparing Specifications

- Bit rate ranges: < 8 kbps - 9.6 Mbps
- Bit widths: 16-24 bits
- Sample rates: 8-192 kHz
- Number of channels: 1-many dozen
- Spectral bins: 128-1024
- Time resolution: 4-12 msec

20

Which Algorithms?



21

MPEG Family

- Moving Pictures Experts Group
- Video and audio compression standards
- MPEG-1, MPEG-2 (MP3), MPEG-4
- Ongoing standardization effort (MPEG-7)

22

MPEG-1 Audio

- 1992
- One or two channels
  - single
  - two independent channels
  - stereo
  - stereo with joint coding
- 32 kHz, 44.1 kHz, 48 kHz
- Specifies bit stream format, decoder structure, but not encoder

MPEG-1 Audio (cont'd)

- Layer 1: simplest; Philips DCC
- Layer 2: more efficient coding; DAB, CD-I
- Layer 3: higher frequency and time resolution; ISDN, Internet
- Decoder for one layer also decodes lower-numbered layers
- Higher-numbered layers have more complex decoder

MPEG-2 Audio

- 1994
- MPEG-2 video used for digital TV, DVD
- Backward compatible with MPEG-1
- Add lower sample rates
  - 16, 22.05, 24 kHz
- 5.1 + up to 7 multilingual/commentary channels
- "MP3" = MPEG-1/2 Layer 3 (not "MPEG-3")

Conclusions

- Entertainment is going digital
  - Audio is a key component
  - Many new market opportunities
    - Internet audio is hot; audio may be the Internet "killer app"
- Audio compression is a key technology
  - Many algorithms, many applications
  - Better algorithms -> better quality, more compression
  - Computation requirements are going up

In the Future…

Photo credit: Dr. Richard O. Duda, SJSU

27

Want Copies of the Slides?

- www.BDTI.com/articles/info_articles.htm

28

Want More Information?

http://www.BDTI.com	**Collection of BDTI's papers on DSP processors, tools, and benchmarking**
http://www.eg3.com/dsp	**Links to other good DSP sites**
comp.dsp	**Usenet group**
Microprocessor Report	**For info on newer DSPs**
***DSP Processor Fundamentals*, BDTI**	**Textbook on DSP processors**

Or, join BDTI... (see www.BDTI.com)

29