The Compression of Everything

Jerry D. Gibson

Department of Electrical and Computer Engineering
University of California, Santa Barbara
gibson@ece.ucsb.edu
What is Compression?

• Whenever we talk on our cell phones, listen to music, view an image, or watch a video, those signals are compressed.
• What does Compression Mean?
  • Broadly it means to reduce the bit rate while retaining quality.
Goal of Data Compression: Reduce Source Bit Rates but Keep Quality

- Source (Speech, Audio, Still Images, Video, EEG, ECG) Compression?
- Goal: Represent a source in digital form with as few bits as possible while still providing an acceptable reproduction of the original
Compression Facilitated the Commercial Success of Key Applications

- 8-to-1 Voice Compression was the key factor in 2G digital cellular
- Still image compression is key to the storage and transmission of images
- Video compression is key to DVDs and streaming of movies and wireless access
- Audio compression is key to MP3
Rate, Bits, and Bandwidth

- Required network bandwidth set by transmitted bit (symbol) rate
- Bit Rate in \( \text{bits/sec} = \text{bits/sample} \times \text{samples/sec} \)
- Sampling rate determined by source bandwidth (Nyquist Rate)—temporal and spatial
Digital Cellular Development

• Analog cellular capacity near saturation in 1988
• An all-digital interface needed a 3 to 6-fold increase in voice capacity (10-fold desired)
• Voice quality and intelligibility needed to be preserved
• Voice compression in early 90’s provided the needed increase—but, it wasn’t easy
Code Excited Linear Prediction

Diagram showing the flow of information through various stages of the code excited linear prediction process, including preprocessing, input speech, LP analysis, quantization, interpolation, synthesis filter, perceptual weighting, pitch analysis, fixed CB search, gain quantization, parameter encoding, transmitted bitstream, and fixed codebook.
How CELP Represents the Voice Signal

![Graph showing formant and harmonic weighting with original speech, formant filter response, and pitch filter response.]
A Few Common Standardized Voice Codecs

• G.711—Narrowband, 64,000 bit/s, 1972
• G.729—Narrowband, 8,000 bits/s, 1996
• G.722.2—AMR-WB/NB, Wideband/Narrowband, 6.6, 8.85, 12.65, 14.25, 15.85, 18.25, 19.85, 23.05, 23.85 kbits/s for WB; 4.75, 5.15, 5.9, 6.7, 7.95, 10.2, 12.2 kbits/s, 2001, 1999
• EVS—Narrowband and Wideband Speech plus Wideband Audio, 2014
Why Compress Still Images?

- Image Size 2048 by 2048 pixels
- Total number of pixels = 4,194,304
- For Color Images, 24 bits per pixel (8 bits each for RGB or YUV)
- Total number of bits to store or transmit for this one image = 100,663,296 bits—100 Million bits
- For 1024 by 1024, this is 25,165,824 bits
JPEG Encoder and Decoder

(a) 8 x 8 blocks

Source Image Data

DCT-Based Encoder

FDCT → Quantizer → Entropy Encoder

Table Specifications

Table Specifications

Compressed Image Data

(b) DCT-Based Decoder

Compressed Image Data

Entropy Decoder → Dequantizer → IDCT

Table Specifications

Table Specifications

Reconstructed Image Data
JPEG Provides

- Rates of 0.25, 0.5, 1, 1.25, 1.5 and 2 bit/s pixel
- 1 bit per pixel/color plane = 3 bits/pixel
- Good quality images
JPEG at 0.25 bit/pixel
Why Compress Video?

• Uncompressed 1080p high definition (HD) video at 24 frames/second
  – Pixels per frame: 1920x1080
  – Bits per pixel: 8 bits x 3 (RGB) = 24 bits/pixel
  – Bit rate: 1.2 Gbits/s
  – 1.5 hours: 806 Gbytes
Why Compress Video--2

• Blu-ray DVD
  – Capacity: 25 GBytes (single layer)
  – Read rate: 36 Mbits/s

• Video Streaming or TV Broadcast
  – 1 Mbits/s to 20 Mbits/s

• Require 30x to 1200x compression
Classical Video Coding

Original value of pixel $i$ in frame $n$

$ f_n^i$

End-to-end distortion of pixel $i$ in frame $n$

$d_n^i$

Encoder-reconstructed value of pixel $i$ in frame $n$

$\hat{f}_n^i$

Decoder-reconstructed value of pixel $i$ in frame $n$ (after error concealment)

$\tilde{f}_n^i$

Intra/Inter

Motion-Compensated Predictor

Motion-Estimator

DeQuantizer/Inv. Transform

Motion-Corrector

Transform/Quantizer

Entropy Coding

Lossy networks

Video Decoder

$ f_n^i$

$ d_n^i$

$\hat{f}_n^i$

$\tilde{f}_n^i$

$ d_n^i$
Comparison of MPEG-2 and H.264 Coding Quality (a) MPEG-2 at 10 Mbps (b) H.264 at 6 Mbps [2]
From [3]

Encoding HEVC

5x - 10x more compute intensive than H.264

HEVC Block Diagram

4K Ultra Definition will multiply compute demands by another 4x - 16x.

Fortunately, many operations can be parallelized.
Speech/Audio Coding Bandwidths
(Uncompressed Bit Rates)

- Narrowband Speech — 200 to 3400 Hz
  (64,000 bits/s)
- Wideband Speech — 50 to 7000 Hz
  (128,000 bits/s)
- Wideband Audio — 20 to 20,000 Hz
  (617,400 bits/s per channel)
Audio Coding Bit Rates

- Uncompressed: 44,100 samples/sec x 14 bits/sample = 617,400 bits/sec per channel
- A 2 minute song = 74.1 million bits or about 9.25 million bytes times 2 for stereo
- MP3 compression at 64,000 bits/sec per channel for an 9 or 10-to-1 reduction
MPEG Layer 3 Encoding and Decoding
Key Steps in MP3 Audio Coding

• Filter Bank Decomposition
• Polyphase Filter Implementations
• Combined Filter Bank/Discrete Transform Implementation
• Decimation/Upsampling Aliasing Cancellation
• Perceptual Weighting
Original ECG data (One Channel)
[Courtesy of Sayood, UNL]
PQRST
EEG 64 Channels [Courtesy of Von Meer, UCSB]
One Filtered Channel of EEG
[Courtesy of Von Meer, UCSB]
Conclusions

• We are highly reliant on voice, audio, still image, and video compression
• We use cell phones, download images and audio, stream videos, watch movies, listen to music on our iPods
• Biological signals are next
• Compression is key to the efficient use of bandwidth and storage
• High quality at the lowest rate possible
• Complexity is not a driving issue!
References

