



MORE TOPICS NEXT...

Speech-Coding Techniques

By:

Joseph Ronald Cañedo

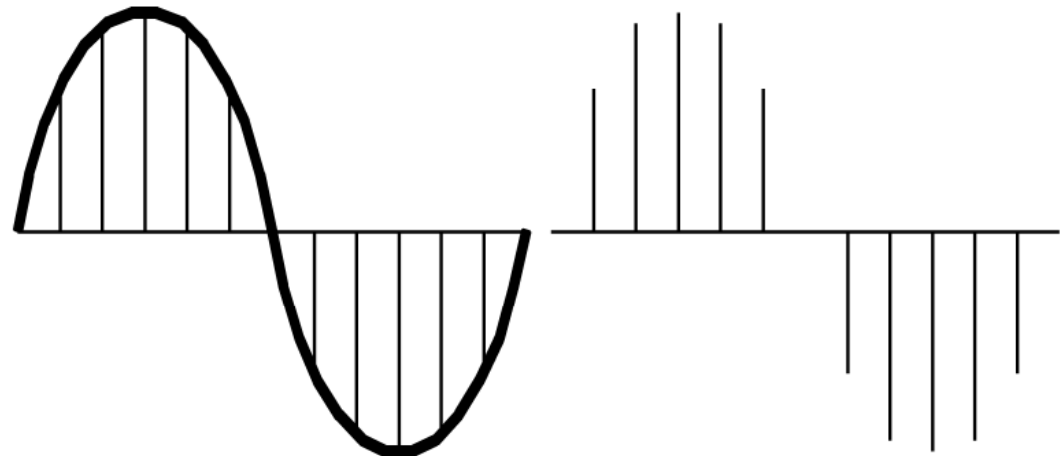
Voice Quality

- The most common speech quality metric is called the Mean Opinion Score (MOS). This measurement scheme is described in ITU-T Recommendation P.800.
- MOS is a five-point scale:
 - Excellent – 5
 - Good – 4
 - Fair – 3
 - Poor – 2
 - Bad – 1

A Little about Speech

- Speech is generated when air is pushed from the lungs past the vocal cords and along the vocal tract.
- The basic vibrations occur at the vocal cords, but the sound is altered by the disposition of the vocal tract, that is, by the position of the tongue or the shape of the mouth.
- The vocal tract can be considered a filter and many codec technologies attempt to model the vocal tract as a filter.

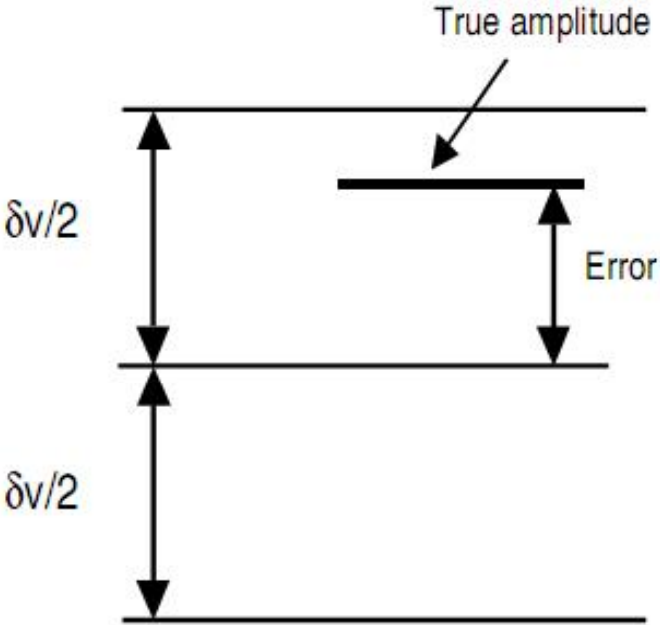
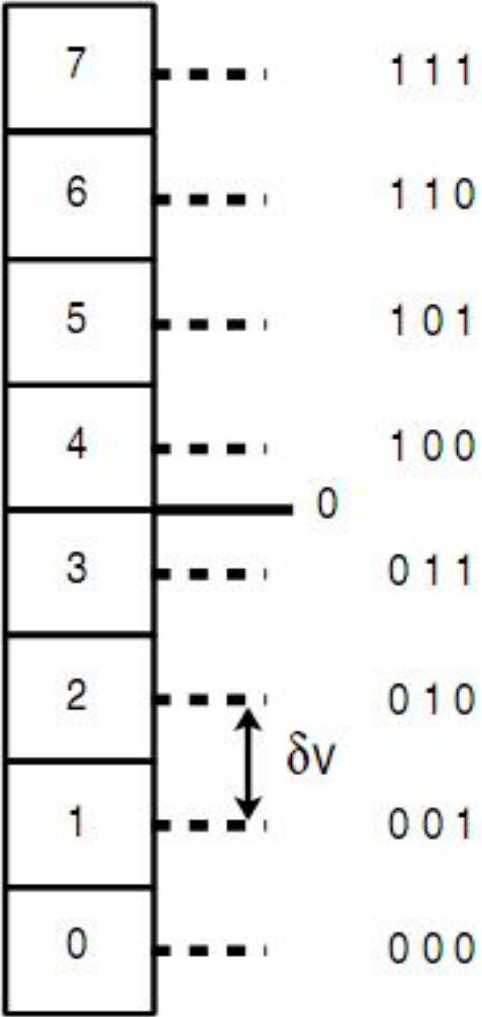
Voice Sampling



- Human Speech Freq:
300 – 3,800Hz
- Nyquist Sampling
Theorem: 8000
samples per second

Quantization

- 3 bits
Quantization



Types of Speech Coders

1. Waveform codecs

- Basically sample and code the incoming analog signal without any thought as to how the signal was generated in the first place.

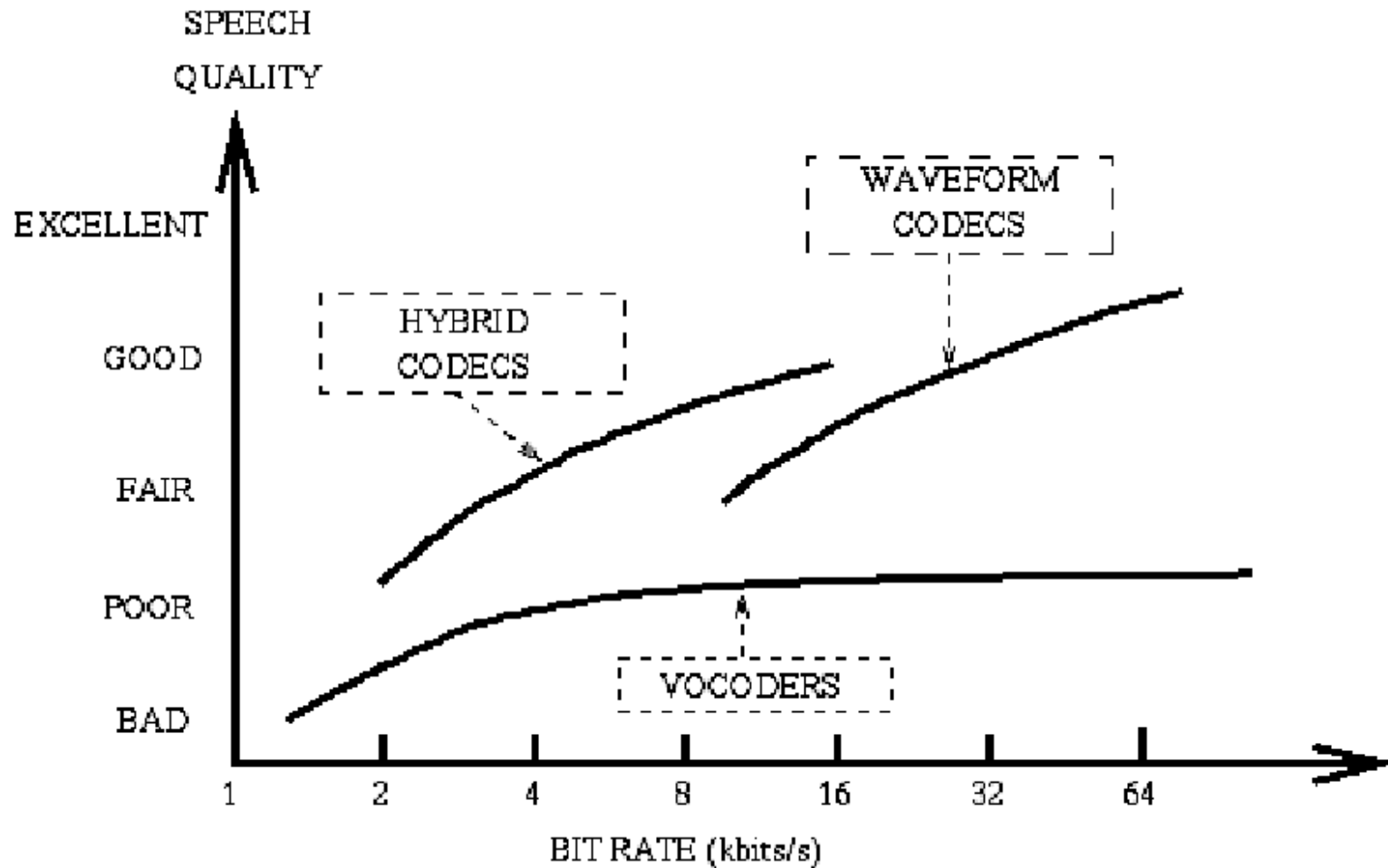
2. Source codecs (aka vocoders)

- Attempt to match the incoming signal to a mathematical model of how speech is produced.
- Operate at low bit rates but tend to produce speech that sounds synthetic.

3. Hybrid codecs

- Attempt to perform a degree of waveform matching, they also utilize knowledge of how people produce sounds in the first place. It tend to provide quite good quality at lower bit rates that waveform coders.

Types of Speech Coders



G.711

- It is a waveform codec and is the coding technique that is used in circuit-switched telephone networks all over the world.
- It has a sampling rate of 8,000Hz, at least 12 bits per sample and a bit rate of 96 kbps (uniform quantization).
- Nonuniform quantization is used, 8 bits used to represent each sample, that leads to 64kbps digital signal 0 (DS0) rate.
- It is often called pulse code modulation (PCM).
- G.711 has 2 variants: A-law and μ -law

Selecting Codecs

- CDMA QCELP, as defined in IS-733
 - This is a variable-rate coder and is currently used on IS-95-based CDMA wireless system.
 - Qualcomm code-excited linear predictor can operate at several rates, but the two most common are the high rate of 13,300 bps and the low rate of 6,200 bps.
- Global System for Mobile Communications (GSM) Enhanced Fullrate, as defined in GSM specification 06.60
- Adaptive Multi-Rate (AMR) codec, as defined in GSM specification 06.90

Tones, Signals, and Dual-Tone Multi-Frequency (DTMF) Digits

- The most sophisticated codecs available today achieve bandwidth efficiency without losing significant quality due to smart algorithms and powerful digital signal processing (DSPs).
- Within the telephone network, in addition to voice, all sorts of tones and beeps are transmitted – not the kind of sounds produced naturally as part of human conversation.
- G.711 can generally handle these tones and beeps quite well.



More tonics next