Enhanced Digital LMR Seminar
19th Aug 2016

Wentworth by the Sea
New Castle NH

LMR Codecs

Why codecs?
Which ones?
Why care?

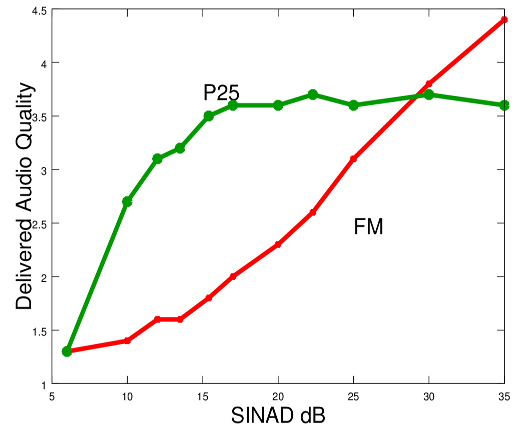
Joseph Rothweiler
Sensicomm LLC
Hudson NH

http://sensicomm.com

Presentation available at:
http://rothweiler.us
Rev. Aug. 22, 2016

Why go Digital?

- **Clearer voice, easier communication (potentially).**
- **Range or Power efficiency**
  - **Longer radio range for the same TX power.**
    - Example: P25 >10dB better than Analog FM.
  - **Resistant to RF adjacent-channel or co-channel interference.**
- **Digital encryption much better than analog scrambling.**
- **Easy to mux multiple audio channels or other nonvoice data (eg, GPS location tracking).**
- **Other features: "voicemail", store and forward, archival recording, etc.**

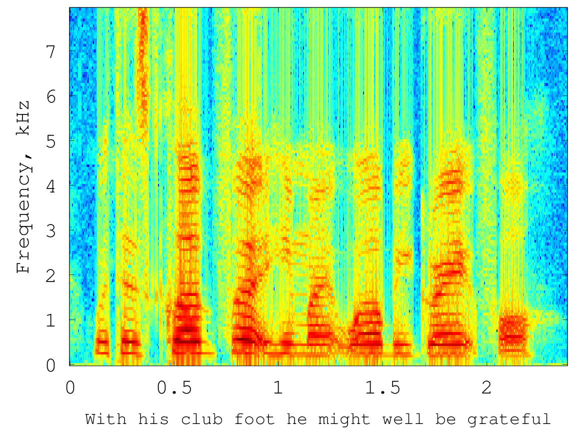


Digital Voice requires about 10dB less transmit power here.

Sensors Signals Comms LLC

Why not use Waveform Coding (PCM)?

- **Waveform coding — Pulse Code Modulation(PCM).**
  - **Microphone → Analog-to-Digital converter; Digital-to-Analog converter → speaker.**
  - **Sample at a fixed rate, with a fixed number of bits per sample.**
- **How fast should we sample?**
  - **Nyquist's Theorem: Sampling rate must be at least 2x the highest frequency.**
  - **Speech goes to 6 to 7 kHz, or more.**
  - **Telephones cut off at 4 kHz.**
- **Number of bits must be high enough to minimize quantization noise.**
  - **16 excellent, 12 good, 8 pretty good with nonlinear (mu-law) coding.**



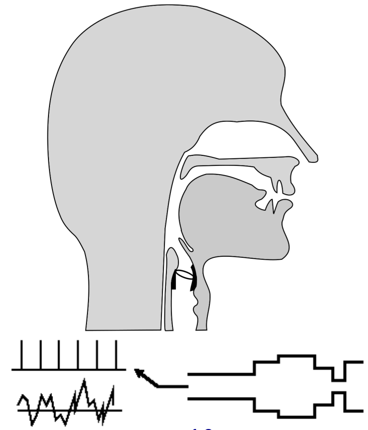
Sampling Rate	Bits	BPS	Quality
16000	16	256k	Near Transparent
8000	8 μ	64k	As good as Bell's 1880's carbon microphones.

- **Simple compression (eg, predictive coding) can get us to 32 or even 16 kbps.**
- **LMR needs 8kbps or below to fix existing channel spacings and bandwidth allocations.**

→ Vocoders to the rescue!

What's a Vocoder?

- **Definitions**
  - **CODEC: Convert analog signal to digital (enCOder) and vice versa (DECoder).**
  - **Vocoder: a codec built on a model of human speech. Vocoders transmit a description of the signal rather than the signal itself.**
  - **The decoder synthesizes a waveform that "sounds like" the input.**
- **Vocoders use an approximation of the human vocal tract:**
  - **Excitation source:**
    - Periodic pulses from the vibrating vocal chords ("voiced"), or
    - noise from friction turbulence ("unvoiced")
  - **Mouth and throat form resonant cavities that shape the spectrum (like a pipe organ).**
- **Simplified mechanical model is a pulse or noise signal feeding cascaded tubes.**
  - **Math equivalent is Linear Prediction model (or autoregressive or all-pole).**
- **Vocoders send**
  - **small set of parameters (filter coefficients, pitch, voicing, gain, ...)**
  - **at a low rate (20-30ms)**
  - **so the bit rate can be very low**



$$y_k = \alpha \cdot x_k + \sum_{p=1}^{10} a_{k-p} y_{k-p}$$

Sensors Signals Comms LLC

A Somewhat Oversimplified Vocoder Grouping

- **LPC: the original buzz-hiss vocoder model.**
  - **SIGSALY(1943) or DoD LPC-10 (1970's).**
  - **Simple 7-parameter model: amplitude, pitch, voiced/unvoiced bit, 10 coefficients.**
  - **Highly intelligible, but very synthetic machine-like quality.**
  - **Problematic in noise, especially the voiced/unvoiced decision.**
  - **Good approach when you desperately need the lowest possible bit rates.**

↓ ↘

- ***MBE* - Multiband-Excited synthesis.**
  - **IMBE, AMBE, AMBE+2, etc.**
  - **P25, satellite systems, etc.**
  - **Replace voiced/unvoiced bit with a multiband mix of buzz and hiss excitations.**
  - **Plus other enhancements.**
  - **Better than classical LPC, without much increase in bit rate.**

- ***ELP* - (something)-Excited Linear Prediction.**
  - **CELP, ACELP, etc.**
  - **TETRA, cellphones, ...**
  - **Use a stored or computed library of excitation functions.**
  - **Search algorithm to select the optimal excitation.**
  - **Can reproduce the input waveform if given enough bits.**

Some LMR Standards and their codecs

Standard	Past, Present and Future Codecs
P25	IMBE, AMBE+2
TETRA	ACELP, AMR 4.75(opt); Future TBD, MELPe?
NXDN	AMBE+2
dPMR	AMBE+2,Chinese TBD, RALCWI, Manufacturer-Specified
TETRAPOL	RPCELP
DMR	AMBE+2
FirstNet	AMR? AMR-WB? Other LTE?
MPT-1327	Analog FM
LTE(cellular)	AMR-NB, AMR-WB, etc
D-STAR(HAM)	AMBE, CODEC2

Some Codec Characteristics

Codec	Year	Voice Rate	Notes
IMBE FR	1995	4.4	TIA-102.BABA; DVSI
AMBE FR	~1998	4.4	TIA-102.BABA; DVSI
AMBE HR	2007	2.45	TIA-102.BABA-1; DVSI
ACELP	1996	4.567	ETSI EN 300 395-2; Reference and test vectors.
MELPe		0.6,1.2,2.4	NATO STANAG 4591
RALCWI	<2011	2.4,2.74	Proprietary cmlmicro.com
RPCELP	1997?	6.0	TETRAPOL specs
CODEC2	?	0.7-3.2	freedv.org - Ham radio project

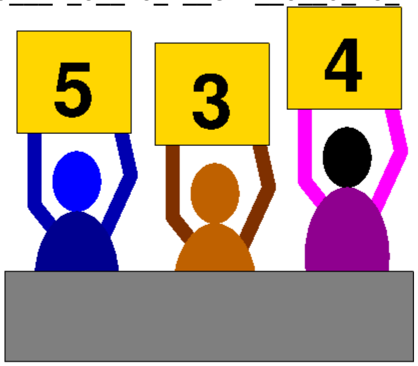
- **New codecs and improved versions of old codecs appear regularly.**

Sensors Signals Comms LLC

A Note on Evaluation

- **Conventional audio measurements (SNR, bandwidth, distortion) don't really apply to vocoders.**
  - **We don't transmit the waveform, so we can't use waveform measurements.**
- **Subjective testing: Rating by a panel of trained listeners.**
  - **Quality: Mean Opinion Score (MOS) - Does it sound good?**
    - Rating as: 5:Excellent 4:good 3:fair 2:poor 1:bad
  - **Intelligibility: Modified Rhyme Test (MRT) Can you distinguish consonants?**
    - Listener hears: "Please select: red" and selects from: led shed red bed fed wed
- **Sounding good is good, but a communication system is for communicating.**
  - **Intelligibility is what you need.**
  - **Tests aren't reality: If the users are complaining, pay attention.**
- **Testing is a really involved process, best left to the experts.**
  - **Data collection, listening environment and equipment,**
  - **listener training, habituation* and fatigue.**

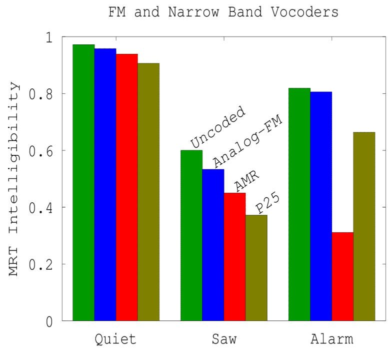
```
In English, consonants carry most of the information.
_n _ngl_sh, c_ns_n_nts c_rry m_st _f th_ _nf_rm_t__n.
I_ E___i__, _o__o_a___ _a___ _o__ o_ __e i__o__a_io_.
```



*"Vocoders sound a lot better on Friday than they do on Monday."

Background Noise Problems

- **Effects**
  - **Theoretical: no longer fits the model.**
  - **Practical: harder to calculate the filter coefficients, and especially the pitch.**
  - **Particular issue for firefighters.**
- **Results**
  - **Lowered intelligibility.**
- **Mitigation**
  - **Avoid noise**
    - Noise-canceling mic, throat mic, dual-microphone setups, etc.
  - **Noise reduction algorithms**
    - Hard to improve intelligibility this way.
- **Robust codec**
  - **Test for it.**
  - **Bit rate tradeoff.**

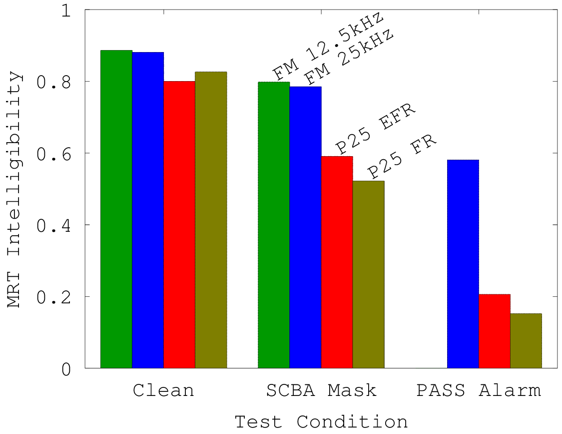


Voice Model Limitations

- Model approximation
  - Nasal sounds (/m/,/n/) don't follow the lossless-tube model.
  - Vocal chord vibration isn't really all that regular for many speakers.
  - Breathy voices: simultaneous mix of voiced and unvoiced excitation.
- Unusual situations
  - Whispered speech (covert situation?)
  - Distorted speech (Firefighter breathing mask - SCBA).
  - Other talkers.
- Result: unnatural sound, difficulty recognizing speakers, loss of inflection (stress, urgency, fear, etc).

Background Sounds

- Background noises may be important for situation awareness.
  - Firefighter alarms: man-down, air tank empty, etc.
  - Sirens, etc.
  - Crowd noise.
- Vocoders are for voice, so nonspeech sounds may be heavily distorted.
  - Which alarm just went off?
  - Is it a happy crowd or a riot?
- Test and tell users about any issues.

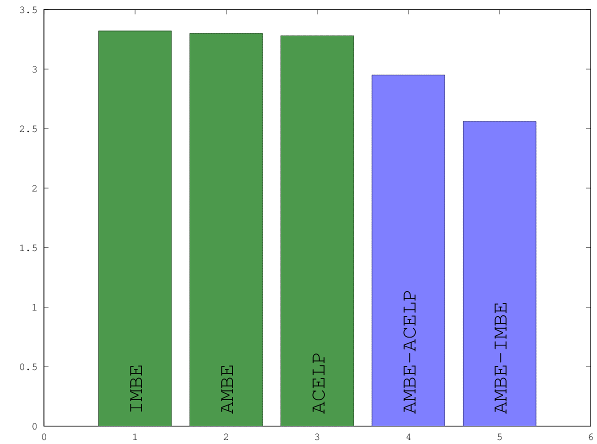


Delay

- Underappreciated issue: Analog FM delay is essentially zero.
- Vocoders analyze 50-100ms segment of audio to extract parameters, so delay is unavoidable.
- Add in processing time, RF synchronization, and error-correction interleaving.
- In full-duplex telephony delays above 500ms become very annoying.
  - Throws off the speak-response rhythm, so perceived as rude behavour by the other speaker.
  - Can't jump in the gap, so end up interrupting while speaking.
- Issue for precise coordination?

Tandeming

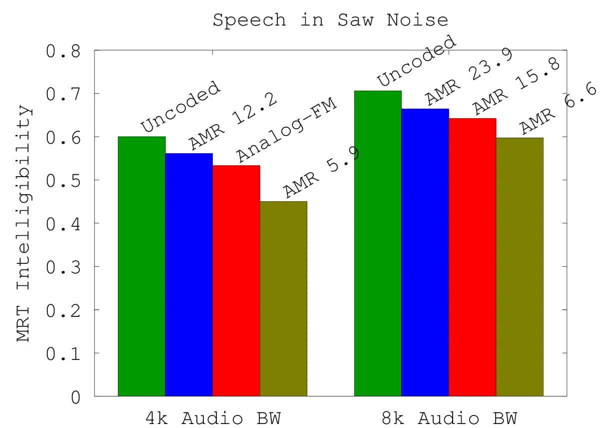
- **Connection of 2 different vocoders in series.**
  - **P25 to TETRA via some sort of gateway.**
  - **Range extension: radio → satellite link → radio.**
  - **Old radio to new radio with improved codec, via gateway?**
- **Avoid tandeming if possible.**
  - **End-to-end transport of codec bitstream is best.**
  - **Consider specifying a mandatory codec for interoperation / backward compatibility?**



Tandem connection is worse then either alone.

Looking to the Future: Wideband

- **Future systems (FirstNet?) may use wider audio bandwidth.**
  - **AMR-WB codecs support 16kHz audio sampling rate.**
  - **DVSI advertises a wideband codec.**
  - **Others exist or are in development.**
- **Wideband does improve operation in challenging acoustic conditions.**
  - **Nice to have when you can get it.**



Bitrate and Audio Bandwidth improve intelligibility.

Summary Recommendations

- **Use the highest bit rate you can get away with.**
- **Test under realistic conditions.**
- **Test for usability (intelligibility), not just "quality".**
- **Think about non-speech signals.**
- **Consider compatibility:**
  - **Need to communicate with other organizations' radios.**
  - **Future buys to expand the system.**

References

- **Images on 3, 8, 9 are from Wikipedia.**
- **Charts on slides 1, 8, 10, 12, 13 were generated by the author from data published by NTIA.**
  - **See these NTIA Publications for the full studies:**
    - http://www.its.bldrdoc.gov/publications/download/TR-01-386.pdf
    - http://www.its.bldrdoc.gov/publications/download/TR-08-453.pdf
    - http://www.its.bldrdoc.gov/publications/download/TR-09-459.pdf
    - http://www.its.bldrdoc.gov/publications/download/TR-13-495.pdf
    - http://www.its.bldrdoc.gov/publications/download/TR-15-520.pdf