Vocoders We Have Loved, Loathed, and Laughed at

DAFX 2009 Keynote

Perry Cook
Princeton University
Computer Science (also Music)
also Sonic Mule

Plan of Attack:

• Pre-History: Vocal Imitation, Shamanism
• Early History: Parlor Tricks, Magic, Ventriloquism
• History: Speaking Machines
• Later History: Communications, Codes and Coding (Information Theory), Cryptography
• Voice as Instrument, Instrument as Voice
• Making Things Sound Like Us: Cross Synthesis
• Making Us Sound Better: Pitch Shift/Correction
All Part of Upcoming Book:

La Bella Voce et la Macchina
(the beautiful voice and the machine)
A History of Technology and the Expressive Voice

Technology: Any human-fashioned tool, technique, method, law, notation, enhancement, etc.

Expressive Voice: Singing, acting, preaching, rapping, praying, etc.

Pre-History: Vocal Imitation
Animal Cries, Magic, Expression

- Imitating birds, animal howls, nature
- Language acquisition itself is/was imitative
- Storytelling, song, music (tonal speech)
- Shamanism, healing, magic, control
- Sternomancy, Ventral Fatiloquency, Gastromancy, Voix et l’ombilique, Engastrimythy, Bauchrednung (Ventriloquism)
- Spirits, divination, séances
- “Spirit Trumpets,” “Spirit Telegraphs,” etc.

A Brief History of /u/a/u/a/ (WahWah)

- Animal Cries => Pre-Lingual Babbling => Hands at Mouth => Object at Mouth
- Lip instruments (brass, mutes)
- Reeds and Free Reeds, and Digeridoo! (vocal tract interactions)
- Electrical Era: Working the tone-control knob (1961, Chet Atkins)
- And finally, the electric Wah-Wah pedal. Right?
- Nope Not quite yet…
Early History: Speaking Machines

"We represent and imitate all articulate sounds and letters, and the voices and notes of beasts and birds"

Francis Bacon (1561-1626) from “The New Atlantis,” 1626

Early Speaking Machines

- Kratzenstein’s (1779) Speaking Tubes
- Abbot Mical’s (1783) Speaking Heads
- Von Kempelen’s (1791) Speaking Machine
More Speaking Machines (-1937)

- Faber’s Euphonia (1840)
- Marage’s Vocal Tracts (1900)
- Riesz’s Vocal Tract (1937)

Speaking Machines: Into The Electrical Era

- Bell: Helping the deaf communicate, Dogs, Cadaver heads, the Telephone, Harp Telephone (Harmonic Telegraph)
- Helmholtz
- Konig
- Stewart (1922)
Dudley’s Voder

- 1939 World’s Fair
- Operators (female) manipulate console

Voder: Source/Filter Model

- Noise/Pulse (wrist bar)
- Pitch control (foot pedal)
- Resonances (10 finger sliders)
- Consonant/Stop Presets (thumb buttons)
Dudley’s Voice Coder (VoCoder)

- Subband decomposition of voice signal
- Voiced/unvoiced source detection/modeling

Later History:
Communications, Coding,
Compression, Cryptography

“Ciphony” (Secure Speech)
Secure Vocoder (US Army/NSA/UK)

- SIGSALY (1940’s)
  10 Bands 0 – 3kHz
  Vacuum tubes
  55 Tons!!
- DEMO!
- Delilah, System X, others
- KY-9 (1953)
  510 lb.
- HY-2 (1961)
- Tom Tremaine

Meanwhile, in the Time Domain…
Linear Predictive Coding (Atal, 1970)

- Waveform: next sample = linear combination of previous samples
- Spectral: best predictor coefficients are least-squares spectral filter
- Perceptual: formant peaks
- Physical: source + ladder filter
- Music composition (Lansky, Dodge, Moorer, Peterson, others 1980’s)
- Speak&Spell, other TI products
- Coders of all types, LPC10, etc.
- to eventually, VCELP (Cell phones)

Parametric speech manipulation: 

Oh No! Did Paul Lansky Invent AutoTune™? !!!!

More on that shortly
LPC for Ciphony

• STU Series (LPC10 secure phones)

STU-III circa 1986

Tremaine

Back to the Frequency Domain

(sort of)
Making Things Talk

- The “Talk Box”
  Sound source excites vocal tract
  Sound source => Hose => Mouth => Mic

- 1976 Frampton Comes Alive, right? Nope…

- 1940 The SonoVox
  Kay Kaiser, others
  Alvino Rey (1939)
  Pete Drake’s “Amazing Talking Steel Guitar” (1964)

- Then Frampton, right? Nope!
- Then, the “VoxWah” pedal
  (1966, Brad Plunkett, Warwick Electronics) demoed by Del Casher

Making Things Talk, The TalkBox Comes of Age

- Finally, the “modern” “Talk Box” (speaker/hose/mic)

- “Frampton Comes Alive” right? Nope, not yet

  Mike Pinera, Iron Butterfly

- THEN Heil’s (1971)
  Stevie Wonder (1972)
  THEN, Frampton (1976)

MOVIES!
The Tragedy of the TalkBox:

Dental Fatigue and Damage

Fortunately there was an alternative…

The Cross-Synthesizing Vocoder

- Vocoder, with non-parametric (audio) source
- Subbands
- Optional /s/
- Optional fuzz
- Optional patching

Harald Bode and Walter Carlos DEMO! “Switched On Bach” (1968) (Clockwork Orange 1971)
The Phase Vocoder

- Lots of bands (hundreds to thousands)
- Track and use magnitude AND phase
- Sinusoidal tracking/resynthesis
- Can implement frequency domain LPC
- Pitch/Time shifting
- Cross-synthesis
- Composition: Dolson, Moorer, others
- Coding/compression: “the burbles and bubbles of the bathtub bells” (Stockham)

DEMO!

Commercial Vocoders (70-80’s)
Korg, Roland, Others

(90’s) PAIA Kit

(00’s) Plug-in Software
Speaking of Commercial Vocoding…

but for a very different purpose

**Autotune™!**

- Antares Inc. (1997)
- Vocoder (LPC+) to
  - Detect pitch
  - Pitch-shift to steer toward target pitch
  - Retune speed [0.0,1.0] sets speed at which correct pitch is approached (0.0 = “Cher”)
- Best selling plug in for last couple of years
- Dirty secret: everyone uses it (not set to 0.0)
- T-Pain and others proudly set it to 0.0
- Newest dirty secret: increasingly used live
Autotune Pre-History

• Rotating head
  pitch/time shifting
  Eltro (1965)

• Eventide Harmonizer (overlap-add, 1976)

• De-glitched Harmonizer (PSOLA, 1977)

• Smart Harmonizer (real/tonal shift, 1980’s)
  AHA! Factor/Agnello etal invented AutoTune?

• Formant correction: Digitech (Studio) Vocalist

Autotune History

• Cher “Believe” (1998)

• Eiffel 65 “Blue (Da Ba Dee)” (1999)

• Soon, countless others

• Soon, everyone (Chicks, Avril, …)

• Soon, many using it live (Reba, Garth, …)

• T-Pain

• then…. the backlash:
  Jay Z “DOA (Death of Autotune)”
  Grammy Blue Ribbon Campaign
Vocoders as Social Commentary

"Autotune the News"

Michael and Andrew Gregory

Take great (and not so great) speeches, moments in history, stupid news, etc. Run them through Autotune. Add music, effects, make videos, put on YouTube, go viral, get famous.

Yea!

So is AutoTune™ Dead?

Tired/Tiring/Cliché?

Annoying?

Become its own Parody? ("Ashlee" MOVIE!)

Industry Outrage/Backlash?

Audience Outrage/Backlash?

Movie!
Probably not

- It’s still a pretty cool “effect”

- T-Pain was top-selling ring-tone last year, and is still (one of) the top Hip-hop artists

- Many want it, many (hip-hop, country, comedy, …) want to do videos with “Pain”

- Worst case, the sound of it will always say “the late 00’s”

AutoTune for the Masses?

T-Pain + Sonic Mule (Smule) + Antares

= “I Am T-Pain” AutoTune™ for the iPhone
Some Conclusions

• The word “vocoder” has (at least) Three definitions:
  – Voice Coder (parameterize human speech)
  – Cross synthesizer (singing cows)
  – Pitch/time manipulator (frequency domain)

The Future?

• Gender/Age Transformation
• Accents (remove/transform)
• Language Interpretation (The Universal Translator)
• Aliens!!
THE END!