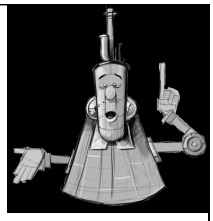
Choirs of the Future?

A look at technology, voice(s) & "cyber choirs (of many forms)



Perry R. Cook, Humbug Sonic Arts
Emeritus Professor, Princeton University Computer Science (also Music)
also Sonic Mule
(and CalArts, ASU, and Stanford CCRMA)

SLEO Workshop
Baton Rouge, LA
April 15, 2012

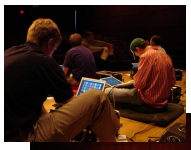
Choir(s) of the Future?

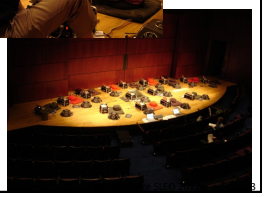
- Fundamental differences between orchestras and choirs, or, more basically, players vs. singers
- Pose and posture, pedagogy, practice, memorization, unions, attire, …
- Cantabile, other "singing" score markings
- Opera, Recit/Aria, Lieder, Art Songs
- Storytelling: The "Song", Oratorio, Theater, Emotion
- Audience/Public Expectations
- Gear: Speakers/mics/monitors/etc.?

Orchestra(s) of the Future?

PLOrk, SLOrk,
and now many other *Lork(ses)

P
S
O
B
Mo
Pho
K
L
V
A
…

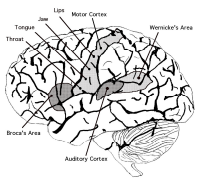


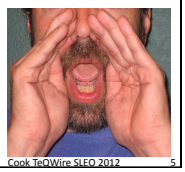


Major Topics:

- History of voices/choirs + technology
- Voice processing (DSP, acoustic)
- Computer analysis of vocal signals (DSP)
- Singer "instrumentation" (measurements, sensors, pose, posture, other physics, & bio signals)
- Voice synthesis, many models (DSP, mechanical)
- Instruments (interfaces) for controlling vocal models (singing machines, computers)
- Voice as controller for "instruments" or other
- Amplification (mic/speaker) issues

Pre-History: Vocal Imitation





Voice: Animal Cries, Magic, Expression, Spiritualism, Coercion

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

Architecture (delay)

Naturally interesting: Canyons, caves

Regular structures

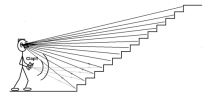
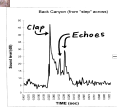


Figure 8.4 Periodic echo pattern from regular stair steps.



Design of space:

- Theaters
- Arenas
- Churches
- Concert Halls

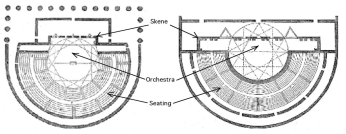


Figure 8.2 Plan of 4th century BCE Greek theater. *Figure 8.3 Plan of typical Roman theater.*

Primitive Voice TeQ

Cupped Hands

Hollow Logs

Megaphones

Early Voice TeQ

"Medicine"!!

Voice as Instrument

- Vocalize (songs without words)
- Humming
- Whistling
- Diddling, Nonsense Vocal Rhythms
- "Reso-Head"
- Clucking, Popping, Other Mouth Sounds
- Ancient and Modern Mouth "interfaces" (seashells, digeridoo, jaw harp, etc.)

Notated (and Instrumental) Voice

Kodaly

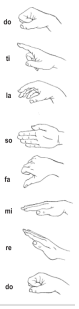
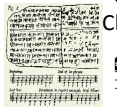
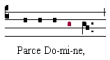


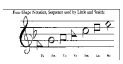
Tabla Bols

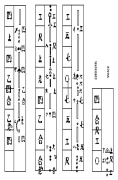
Voice-specific Notations
Instrument-linked
Vocal/Visual/Gestural
Hand-based Notations
Conducting
Clix, other sync.

Guido







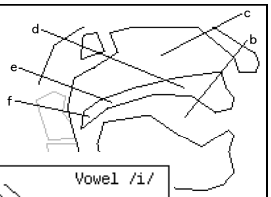


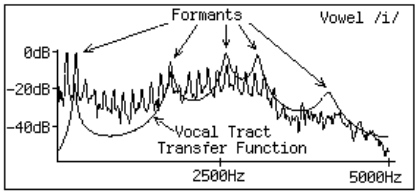
Shamisen

Earliest Greek Gregorian Shape Note

The Voice

Source/Filter Models
(articulatory models)





Spectral Models (perceptual models)

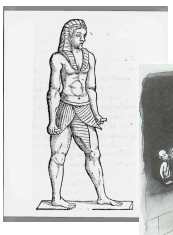
Early "Synth" History: Speaking Machines

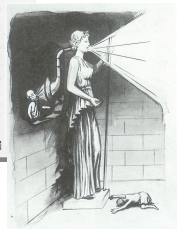
"We represent and imitate all articulate sounds and letters, and the voices and notes of beasts and birds"

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

Mechanical Voices

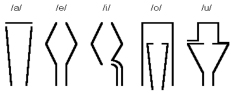
- Ancient Speaking Statues
- Other Voices from (pre-formal) History
- Acoustic Tubes
  - Kratzenstein
  - Paget
  - Marage
  - von Kempelen
  - Faber
  - Reisz
  - Wheatstone
  - the Bell Family
- Spectral Models (Helmholtz, Koenig)
- Waveform Models (Konig, Preece and Stroh (Phonautograph too))





Early Speaking Machines

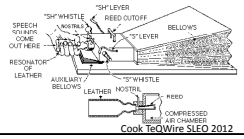
- Kratzenstein's (1779)



(vox humana)

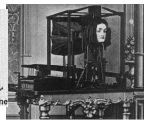
Speaking Tubes

- Abbot Mical's (1783) Speaking Heads
- Von Kempelen's (1791) Speaking Machine

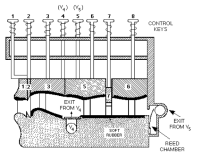


More Speaking Machines (-1937)

Faber's Euphonia (1840)



Marage's Vocal Tracts (1900)



Riesz's Vocal Tract (1937)

Speaking Machines: Into The Electrical Era

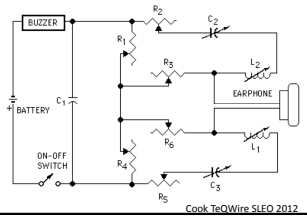
Bell: Helping the deaf communicate,
Dogs, Cadaver heads, the Telephone,
Harp Telephone (Harmonic Telegraph)

Edison

Helmholtz

Konig

Stewart (1922)



Dudley's Voder

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

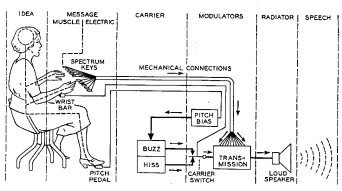
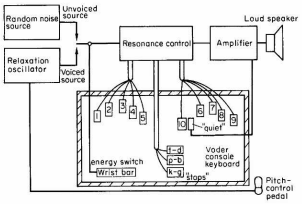


Fig. 8—Schematic circuit of the voder.

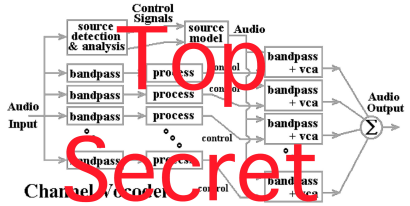
Voder: Source/Filter Model

- 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



Top Secret

- Immediately CLASSIFIED by US Government

Making Instruments Talk

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:
  Harmonica, Flute, Sax
- Electrical Era: Working the tone-control knob (1961, Chet Atkins)
- And finally, the electric Wah-Wah pedal. Right?
- Nope (another story)

Making Instruments 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 Steel Guitar" (1964)
- Then Frampton, right? Nope!
- Then, the "VoxWah" pedal (1966, Brad Plunkett, Electronics) demoed by Del Casher

Talking

Making Things Talk, The TalkBox Comes of Age (1969+)

- Finally, the "modern" Talk Box (speaker/hose/mic)
- "Frampton Comes Alive" right? Nope, not yet
- "The Bag", Kustom Electronics (Doug Forbes, 1969)

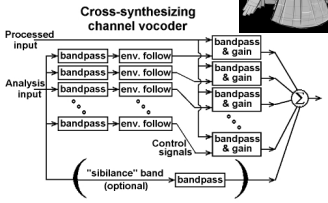
Mike Pinera, Iron Butterfly

THEN Heil's (1971)
Stevie Wonder (1972)
Jeff Beck (1975)
THEN, Frampton (1976)

The Cross-Synthesizing Vocoder

- Vocoder, but alternate audio source
- Subbands
- Optional /s/
- Optional fuzz
- Optional

Harald Bode



and Wendy Carlos
"Switched On Bach" (1968)
(Clockwork Orange 1971)

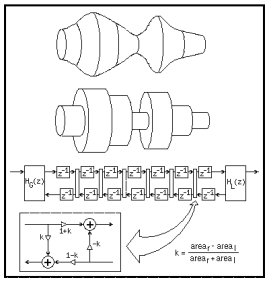
DEMO!?

Speaking/Singing Machines and Synthesizers (cont.)

- Acoustic Tubes (von Kempelen, Kelly-Lochbaum)
- Vocoders/Voders (Dudley, Bode/Carlos)
- Variable Filter-based Models (Klatt, Sundberg)
- Linear Prediction (Atal, Dodge, Lansky)
- Formant Wave Functions (FOFs) (Rodet, Jaffe)
- Sinusoidal Models (McAulay/Quatieri Serra/Smith)
- Frequency Modulation (Chowning)
- Acoustic Tubes (Kelly-Lochbaum, Cook, Carre, Liu, Kim)
- Template-based models (Lomax, George, UPF)

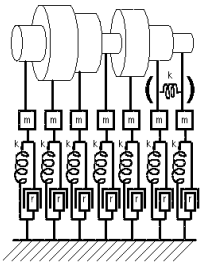
Acoustic Tubes

- Mechanical
- Analog Electronic
- Digital Articulatory Ladder Filter
  Bell Labs 1961
  (Kelly, Lochbaum & Mathews)



Acoustic Tubes (improved)

- Acoustic Tubes (Cook, Carre, Liu, Kim)



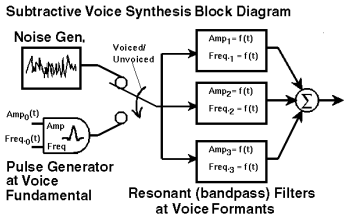
- Nasal Tract
- Throat Radiation
- Lots of Real-Time Control
- Shape/Spectrum Feedback
- Articulation Tracking
- Inertial Articulation Modeling

Greek Shiela

Daisy Duet

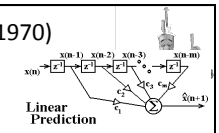
Source/Filter Models

- Linear Prediction (Atal, Dodge, Lansky)
- Formant Filter Models (Fant, Rabiner, Klatt, Sundberg)

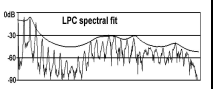


Choir (Ternstrom 89)

Linear Predictive Coding (Atal, 1970)

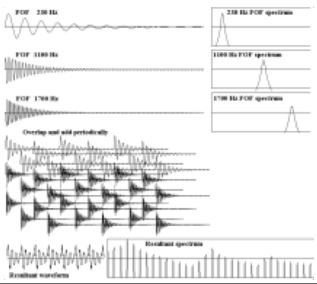


- Waveform: next sample = linear combination of previous samples
- Spectral: best predictor coefficients are least-squares spectral filter
- Perceptual: formant peaks
- Physical: source + ladder filter
- Cross-synthesis
- Music composition (Lansky, Dodge, Moorer, Peterson, others 1980's)
- Speak&Spell,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



Formant Wave Functions

- Formant Wave Functions (FOFs) (Rodet)
  Formants in the time domain

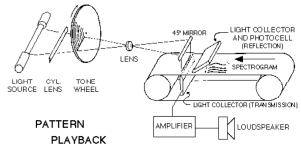


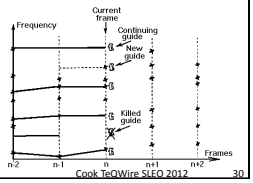
Impossible Animals (Jaffe)

Gesualdo (Rodet/ Bennett)

Spectral Models

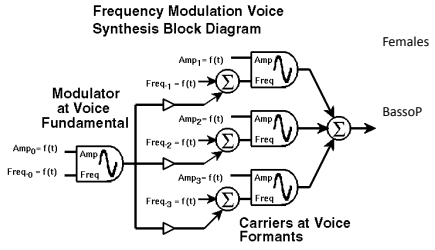
- Helmholz, Konig
- Haskens Labs PAT (1949)
- Sinusoidal Models (McAulay/Quatieri 1986)
- Sinusoidal Models + Residual (Serra/Smith 1989)
- Sines, Residual + Transients (Verma/Meng 1998)





Frequency Modulation

- Frequency Modulation (Chowning)



Template-based Models

Lyracos (Lomax, George, UPF)

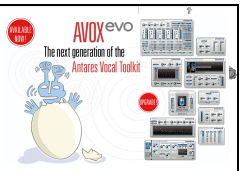
LomaxElvis

LomaxElla

Vocaloid (UPF, Yamaha)

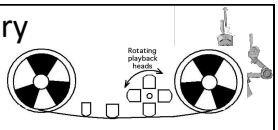
And…..
We can't talk about voice, technology, and synthesized singing these days without…

Autotune™!



- Antares Inc. (1997)
  - 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 (1965) (motivates cheap digital pitch shifter)

DEMO

- 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

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!

AutoTune for the Masses?

T-Pain + Sonic Mule (Smule) + Antares

= "I Am T-Pain"
AutoTune™ for the iPhone

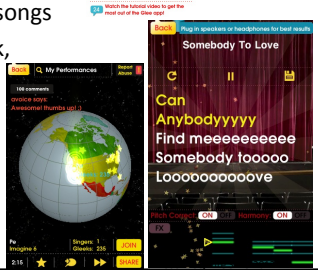
OK, Enough AutoTune

Let's talk about how mobile and ubiquitous technologies have been, and can be used for choral purposes

SMule: Social Music

Glee!

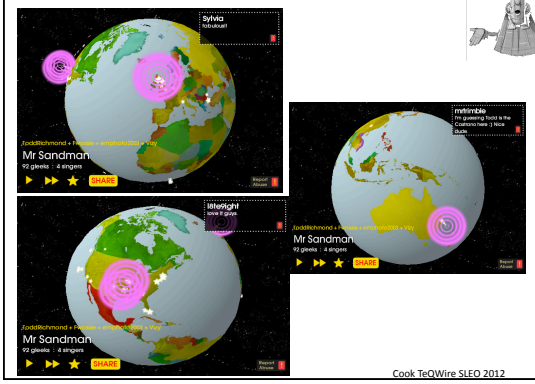
- Pitch Correction and Harmonies
- Songs from the Show
- Other (non show) songs
- Share via Facebook, Smule, Email
- Global "Glee Clubs"
- Ad Hoc Groups
- Earn "Gleeks" Comments



People who know each other,

but geographically separated, singing together

Strangers making ensembles



Massive (3000 and growing)

- Bicycle Built for Two Thousand (2009) (Aaron Koblin and Daniel Massey) uses Amazon Mechanical Turk
  - Present a audio token "snip"
  - User records imitation of that
  - Do this to cover the whole song multiple times
  - Collect and mix all
  - Animate with "piano-roll" style display

- Lux Arumque, Eric Whitacre (2010)
- YouTube call
- YouTube submissions
- Assemble video/audio
- Make cool "world choir" video

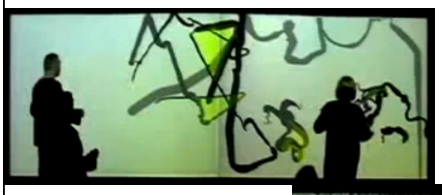
Liberteria, Sabrina Peña Young (2013)

- Virtual Animated Opera
- Casting calls via Facebook, other
- Auditions by submitted recordings
- Click/backing tracks on BandCamp
- Vocal parts submitted by email/upload
- Editing/mixing by distributed team
- Multiple singers per role, for morphing
- Animation by distributed team
- Sound/Score driven facial animation
- Many cast/composer have never met
- Screen in 100+ locations in 2013

Speaking of Animation, and Video…

- Video unrelated to performers
- Performers controlling video
- Video tracking (faces, bodies)
- Display to performers (scores, other)
- Remote (Telematic) performances

Messa di Voce (Levin 2003)



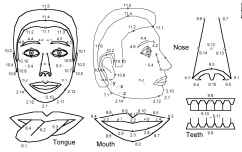
- Camera tracking +
- Voice Analysis +
- Real-time Animation

000000swan

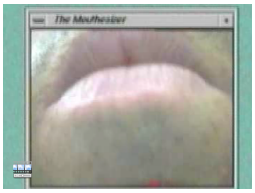
- Kinect tracking
- Sensors, mics, instruments (cello)
- Video display
- Goth/tribal costumes
- Wekinator
- Video(s)

Mouth/Face tracking (body too)

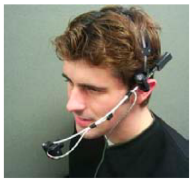
- Very Nervous System
- EyesWeb (VAChorale movie)
- Processing Video Demo
- Jitter
- FaceOSC
- MAGE (vid)
- Mouthesizer



The Mouthesizer (Lyons, ATR, 2001)

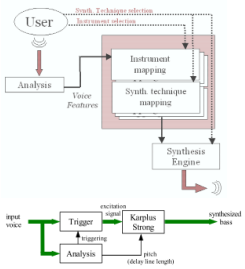


"Boom" Camera + Image Processing

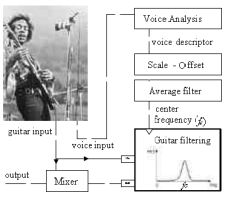


Voice-Controlled Synthesis and Processing (2005, UPF)

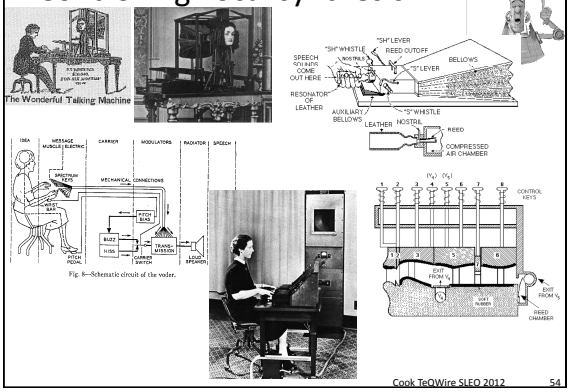
Voice-controlled Bass and Synth



WahWactor: Voice-controlled WahWah



Controlling vocal synthesis

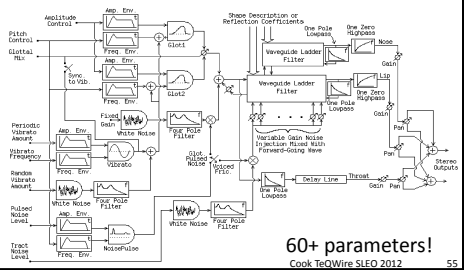


Controlling Vocal Synthesis

- Real time, so possible
- But many, many parameters
- Not a natural "fit"

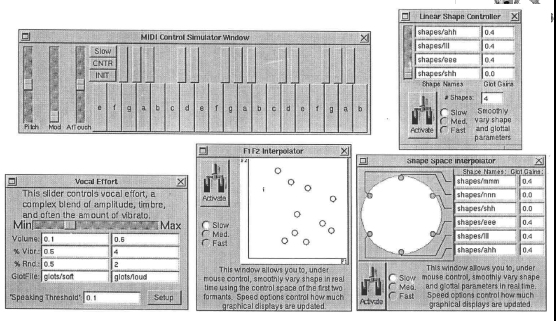
Uh….

Huh?



60+ parameters!

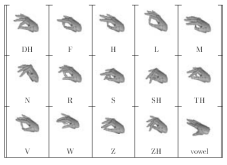
Few-to-Many Controls ('93)

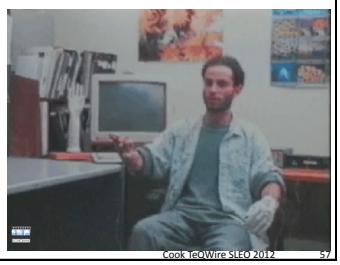


Tract/Glot, Linear, F1F2, Arbitrary 2D

GloveTalk (Fels and Hinton, 1990+)

- Data gloves and 3D position to control speech synthesizer
- Lef Hand "Macros" for Consonants
- Pitch height
- Vowel Space





SqueezeVox (with Colby Leider 2001)

- **Voice Control Issues:**
- **Pitch**
- **Breathing**
- **Articulation**
- **etc.**
- **Fix: Accordiae?**

Lisa

Maggie

Colby's Bart

The COWE

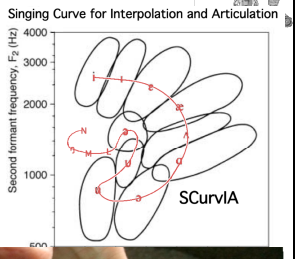
Controller, One With Everything

"Bagpipe" with Tilt Sensors, Linear FSR, Buttons, and more

A New One-to-Many Map:

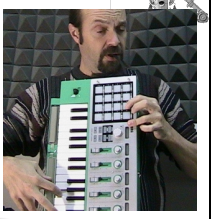
SCurvIA (2011)
Singing(Synthesis) Curve for Interpolation of Articulation

"Rainbow" vowels and liquid consonants arranged along a single line



2004: More Voice Control

Voice
Oriented
Melodica
Interface
Device

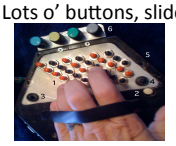


VOMID

Blowing, Keyboard, Linear FSR, 3xTilt, Joystick, Buttons, Knobs, Sliders, …

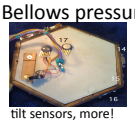
Re-Design: Maggie 2009

Lots o' buttons, sliders (more buttons)



And yes, Batteries

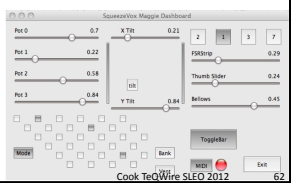
Bellows pressure



AND! Wireless

tilt sensors, more!

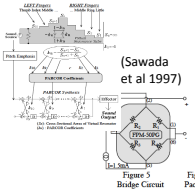
GUI Dashboard and Proxy



Others

Sounds in Hands

Phonodeon (2005)
(Anastasia Georgaki)

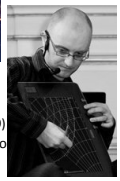


(Sawada et al 1997)

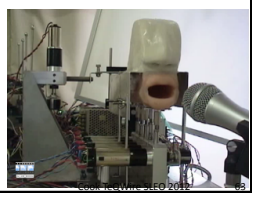
Figure 5 Bridge Circuit

Figure 6 GraspMIDI: 4 P-Units Packed into Spherical Silicon Ball

Talking and Singing Robot (Kagawa University)



HandSketch (2009)
(Nicolas d'Alessandro

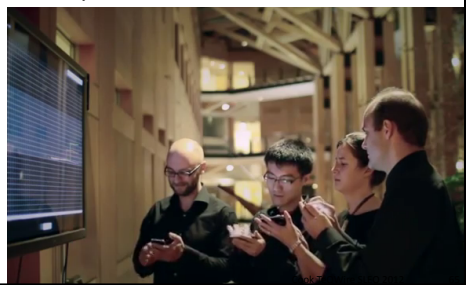


GRASSP/DIVA (UBC)

- Gesturally-Realized Audio, Speech and Song Performance
- DIgital Ventriloquized Actors
- Extension of GloveTalk Projects

Choir Mob (Astrinaki and dAllesandro 2011)

- iPhone Singing Synthesis App
- "Intertwine" by Aura Pon, for Choir Mob and Vuzik



Choirs of the Future?

- Players/Dancers/Actors Controlling Digital/Mechanical Vocal Models?
- Robots? Androids? Cyborgs? DIVAs? Toys?
- Computers Singing to Other Computers?
- Vocaloids Judging AutoTuned Humans? (CyberIdol™ on FOX!)
- Hey! How About Augmenting Singers?
  (perhaps the optimal vocal controller…)
- draw experience from…

Orchestra(s) of the Future!

Future (past, now) "Cyber" Singers

- Ursula Dudziak 1970's+
- Laurie Anderson 1980's+ (+Video)
- Pamela Z. 1990's+
- Amy X. Neuburg 1990's+
- Bjork! (rhymes with PLOrk)
- Many others
  what about choirs??)
- Many works for "choir and tape" (70's)
  Stockhausen, Tavener, Pinkham, Schafer, Kemner, Winsor, Adams, ...

Choirs of the Future

- Effects added to singer voices
  (controlled by singers)
- Synthesis via controllers
  (controlled by singers
  dancers, audience?)
- Integrating this + more

Courses: TeQWire: Voice Technology Seminars
Two-Sided Plays (with Laurie Anderson)

Choirs of the Future: TeQ

Singer signals, and sensors to measure:

- Pose, Posture, Joint Angles
- Head, Face, Mouth
- Bio Signals
- Pitch, Loudness, Timbre, other acoustic

DSP to do:

- Processing of live singing, canned singing, pre-recorded sounds
- Feature Extraction and Machine Learning

Bio Signals and Sensors

- Breathing (events, depth, rate)
  (mask, hot-wire, fan, diaphragm, bernoulli, chest band, cond/resistive fabric)
- EGG (Ω between vocal folds, neck band electrodes)
- EMG (electrical muscle signals, skin electrodes)
- EKG (heart EMG, electrodes !!CAUTION!! pulse, audio)
- EOG (electrical/magnetic eye direction, coils)
- EEG (brain: α, β, …, electrodes, coils, squid, IR)
- GSR (skin resistance, electrodes, pennies ☺)
- Pulse (heartrate, EKG, electrodes, pulse oximeter)
- Blood gasses (O_2, N_2, CO_N, … IR, other) (VAC vid)

Physical (Pose) Sensors

- Tilt (head, hands, torso, foot/feet, …, accelerometers, video)
- Absolute spatial positions/orientations (6DOF) (body in room, head, hands, feet, etc., RF, polhemus, video, IR, tethers!) (DEMO)
- Joint angles (arms, wrists, neck, fingers, hands knees, ankles, torso, cuffs, bendies, gloves, video)
- Facial: orientation(s), mouth/lips, eyebrows, eyes, gaze/focus, video, IR with markers) (VAC vids)

Recent solutions for some of these: Kinect! , iPhone! (DEMO)

DSP Things To Do

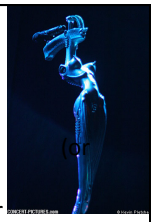
- Delay, loops, chorus, flanging, …
- Pitch shift (correct), harmony
- Spatial processing
- Gender/character (accent?) modification
- Sinusoidal resynthesis (with modification)
- LPC/Vocoder modification/cross-synthesis
- Playback of pre-recorded (contributed) audio
- Gating of audio (captured or pre-recorded)
- Scrubbing of audio (captured or pre-recorded)

Audio Features to Extract
(and things to do with them)

- Time Domain:
  - Loudness (peaks, power)
  - Zero Crossings (Noisiness)
  - LPC coefficients (cepstrum)
- Frequency Domain:
  - Spectral Flux, Centroid, Tilt, Rolloff(s), Sub-band energies, Formants, Harmonicity, Harmonics to Noise Ratio (HNR), Pitch, Multi-pitch
- Classification/MIR/Machine Learning:
  - Vowel/Consonant, Phoneme, Gender, Singer ID, …
- SMIRK, CLAM, Weka, Wekinator

DEMOs

Microphones/Amplification



- Mic Stand(s)
  (Freddie, Korn, Emic (NIME03), grunge (2mics))
  How about sensors on/in mic stand?
- Microphone in/on music stand concealed in folder, or in the laptop?)
- Sensors in/on music stand, or in folder
- Head(set) microphones
  Countryman, Shure, Sennheiser, Audio Technica)
- Wireless mics, mic/instr. packs, in-ear monitors
- Wireless sensor systems (MIDI(Tron), Spark, Arduino)
- Speakers: House sound? Monitors? Local? Wearable? (remember precedence)

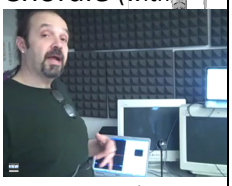
Haptics, Robots, ???

Haptics (display to singers, audience?)
Robots (with singers, as singers)
Exoskeletons, prostheses, …
Waisvisz (the Hands), BugMudra, Stelarc
Karmetik Machine Orchestra

The Virtual Augmented Chorale (with

Ben Knapp, 2005-2009)
Part of the
"Integral Music Controller"
Signals/Inputs:

- Physical/Gestural (pressure, tilt, etc.)
- Bio (EMG, EKG, EOG, O_2)
  (VAChorale music stand)
- Video: Face, Body
- Emotion (inferred from bio signals listed above)
- Other



Breathing => Launch phrases

Arm up/down => #singers

Left grip => scrub mode

Right tilt => rate

SSSS: Singer's Sensor Speaker Stand

Some Recent CyberVoice Pieces I've Done

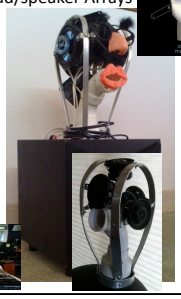
- CredoI (solo – trio) (2010)
- I Want (solo) (2012)
- LOrX Aeterna (PLOrk, April 7, 2012)

NogginSonics ("head art")

Head-Shaped Speakers/other

Spelmet

Head/speaker Arrays



PhoNoggin (ipa)

CraniOrgan

SolFoggins (pitch)

Conclusions

- Our voices are primal
- Humans perceive the world through (nearly) hard-wired linguistic "lenses"
- Singing is synonymous with "beauty"
- Synthesis methods abound
- Sensors for new controllers abound
- DSP and machine algorithms abound
- Much research remains to put them together
- Figuring out what to do is the hard part
- Magic (aesthetic transcendence?) is possible
- Goal: Make Art!!
- Cantibile (perform in a singing-like manner)

Open Questions, Vexing Problems

- The Curse of Causality
- "Grey Paint"
- "What the hell are they doing?" (a common issue)
- Robustness (or at least less fragility)
- Repeatability, reliability, flexibility
- Persistence (of "instruments", pieces)
- Feedback!!!

Things I Ain't Seen (much of) Yet

- Audience involvement
- Masks (Greek, dance/theater, puppets, but little in the way of cool EFX in masks)
- Features and Machine Learning on whole choirs (bio signals, acoustic signals)
- Singer/Choir-specific speakers
- Integrated wireless systems (mic, and sensors, and in-ear monitors)
- Actual dedicated cyber-choral ensembles
- Vocoder choir, sonovox choir, speaking machine choir…
- Theremin-like instruments, ensembles, audience theremin
- Tradition: pedagogy, practice, repertoire, (yea, sure)

Resources

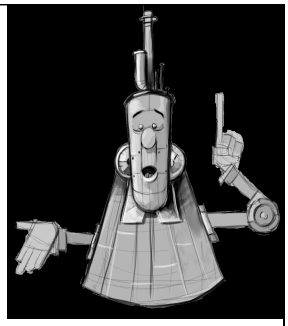
- Gear (sensors, electronics):
  - ElectroTap
  - Infusion Systems
  - Sparkfun, AdaFruit, Parallax, Arduino.cc, AllCorp, Jameco, DigiKey, others.
- Speakers: ElectroTap, Isobel, Bluetooth, ???
- Software:
  - Max/MSP, PD, Processing, SuperCollider, Aura/Nyquist, C*(monkeys), STK, ChucK, Java
- This talk: vocebella.org

ACKs

National Science Foundation
Guggenheim Foundation
MacArthur Foundation
Ariel Foundation
David A. Gardner '69 Magic Fund
(Princeton Humanities Council)
Kimberly and Frank H. Moss '71
Research Innovation Fund
(Princeton SEAS).

NIME, PLOrk, TeQWire, LAP, Colby Leider, Dan Trueman, Ajay Kapur, Ben Knapp, Curtis Bahn, Ge Wang, Rebecca Fiebrink, and the Princeton CS and Music Graduate Students.

Thank You!



Drawn from Upcoming Book:

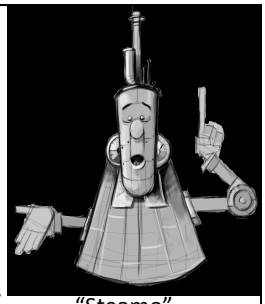
La Bella Voce e la Macchina

(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.



"Steamo"

(fictional turn of century steam-powered opera singer)

Table of Contents (Section 1)

Contents (Section 2, Technology)

Contents (Section 3, Culture)

Contents (Section 4, Machines)

Contents (Appendices)

Interviews and Supporting Materials: