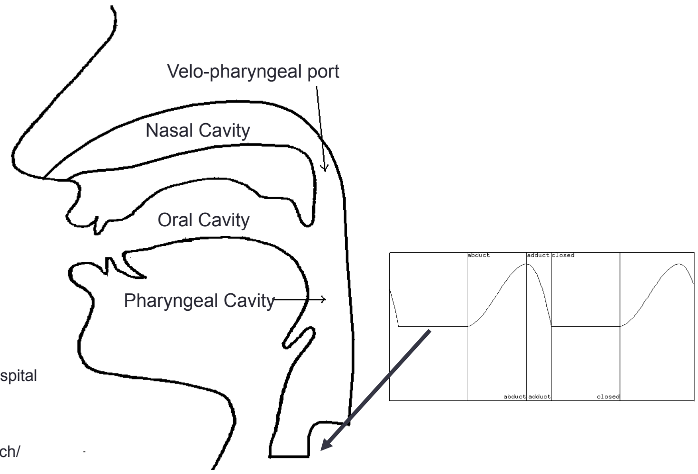
INTRODUCTION TO COMPUTER MUSIC
THE HUMAN VOICE

Roger B. Dannenberg
Professor of Computer Science, Art, and Music

INTRODUCTION

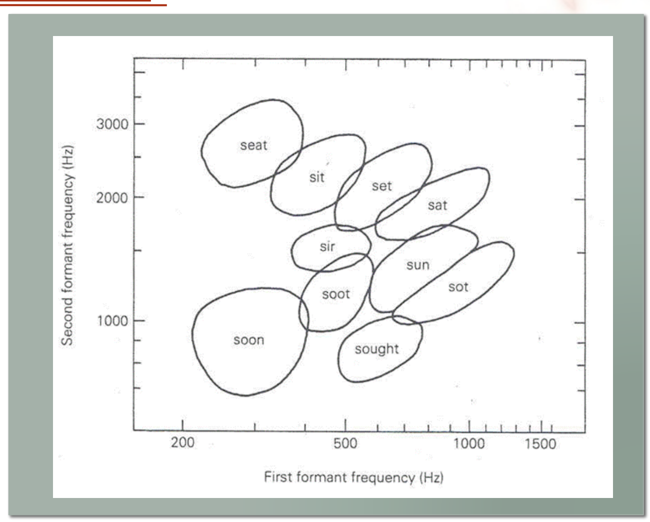
Source Filter Models

How The Voice Works



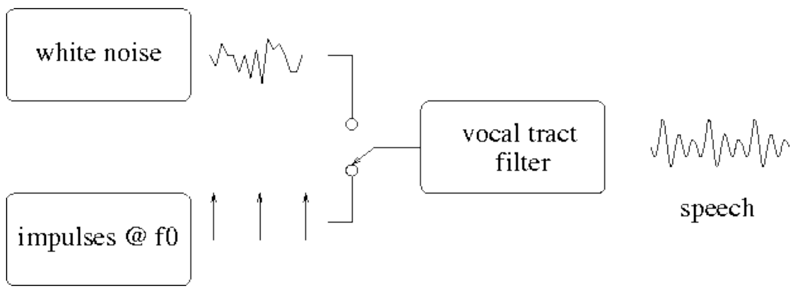
Images from: duPont Hospital for Children and the University of Delaware, *Speech Tutorials*, http://www.asel.udel.edu/speech/tutorials

Formants and Vowels



http://www.phy.davidson.edu/fachome/dmb/digitalspeech/formantmap.gif

Source-Filter Model



From Tony Robinson, *Speech Analysis,* http://svr-www.eng.cam.ac.uk/~ajr/SA95/SpeechAnalysis.html

LPC

A Speech Analysis/Synthesis Method

LPC

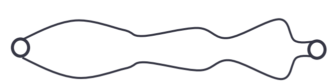
- LPC = Linear Prediction Coding
- Model: predict next sample as a weighted sum of past samples.

$$s_n = \sum_{i=1}^{p} a_i s_{n-i}$$

- This formulation gives rise to an *allpole* filter: the response consists of resonant peaks.
- LPC analysis finds the filter with that best approximates the signal spectrum.

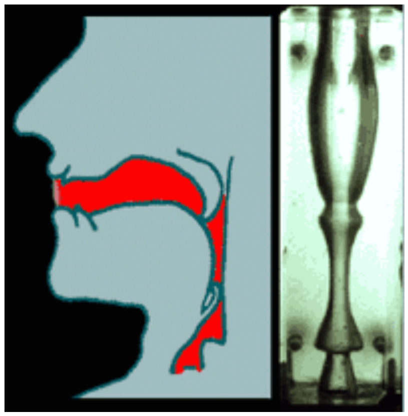
LPC Analysis

- The physical analogy is a tube with varying cross-section:
- Conducted in frames (analogous to short-time windows in SFFT)



- Frames give rise to changing coefficients, which model changes in tube geometry (or vocal tract shape)

Acoustic Tube Producing "AH"

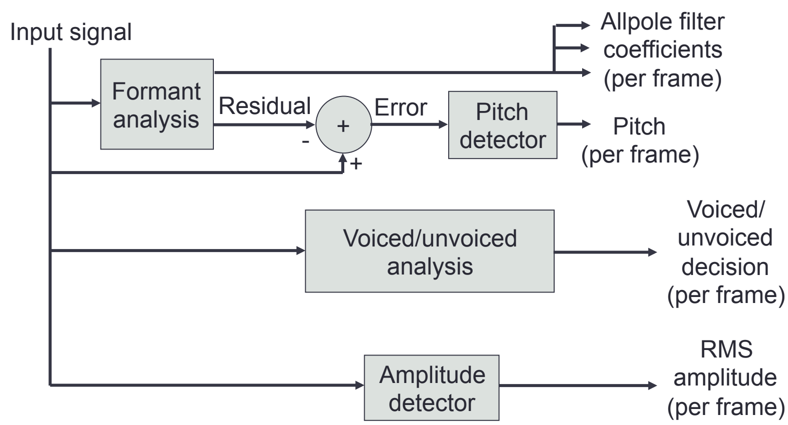


From: the Exploratorium. http://www.exploratorium.edu/exhibits/vocal_vowels/vocal_vowels.html

LPC Analysis, continued

- LPC creates an inverse filter.
- Applying inverse filter gives a *residual.*
- Residual may either be an estimate of glottal pulses → do pitch analysis to estimate source
- Or noise → use noise model for source

LPC



Musical Applications

- Replace source with some other sound
- "Warp" the filter frequencies
- Modify the source and LPC coefficients (glottal pulses or noise) to perform time stretching
- See demos/lpcdemo.lsp

VOSIM

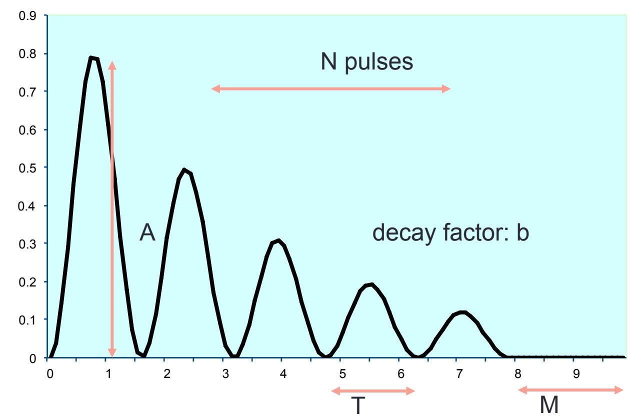
A simple and fun synthesis method inspired by the voice

VOSIM

- Voice-inspired technique.
- Developed in 70's by Kaegi and Tempelaars
- What happens when a glottal pulse hits a resonance?
  - Answer: exponentially damped sinusoid

VOSIM Parameters

- VOSIM uses a pulse train of sin^2 pulses.



VOSIM Application

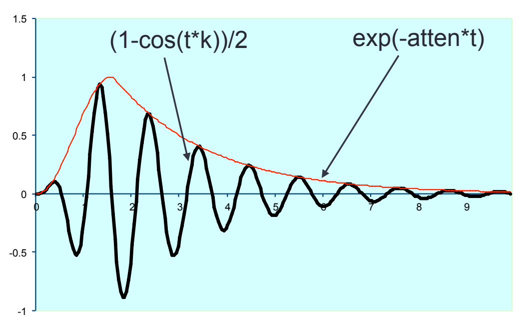
- One formant (resonance) per VOSIM oscillator
- T gives formant position
- M + NT is period
- Tempelaars used various "delta" or "increment" parameters to get change over time
- See vosim.sal example code
- Some sounds on youtube: `https://www.youtube.com/watch?v=7GetTjx96D0`

OTHER VOICE SYNTHESIS TECHNIQUES

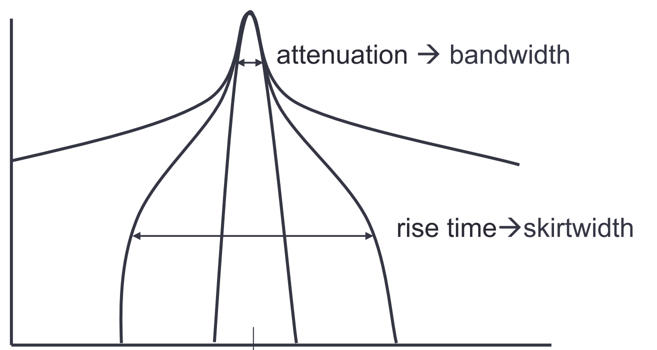
FOF, Vocoder, MQ, SMS

FOF

- FOF is a French acronym "Formant Wave-Function" (Synthesis)
- FOF simulates the effect of a glottal pulse exciting a single formant resonance
- Related to VOSIM



FOF in the Frequency Domain

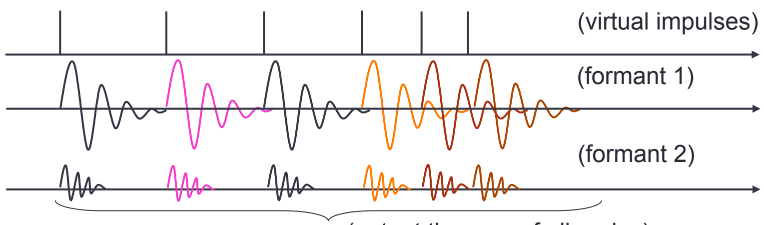


FOF Analysis

- FOF parameters can be obtained automatically by modeling the peaks of an STFT
- Not limited to voice synthesis

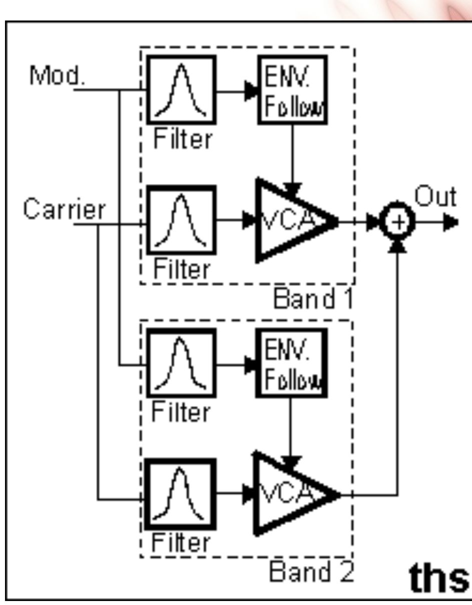
FOF Synthesis

- Each FOF generator creates the response of a single formant to a single impulse -- essentially a grain.
- FOF generators are allocated dynamically (since response to one impulse can overlap in time with response to the next impulse)



Example: Jean-Baptiste Barriere, *Chreode I*, `https://www.youtube.com/watch?v=5AEFhybYrPg#t=128`

Vocoder



From: http://www.ths-nation.de/recall/vocoder.htm

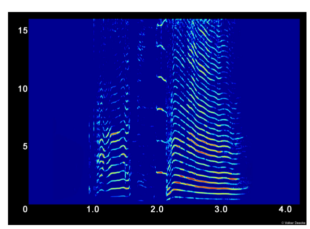
Phase Vocoder

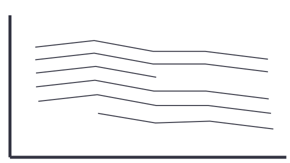
- Intended mainly for sounds with few partials
- Assume each partial lands in a separate Discrete FT bin/ channel/frequency
- From successive frames of FFT, estimate amplitude and phase of partial
- Reconstruct using inverse FFT

Original Stretched Lower

MacAulay-Quatieri (MQ) Synthesis

- Developed as a speech compression technique
- Identify and track sinusoids in STFT frames
- Representation is list of trajectories of sinusoidal partials
- Additive synthesis





MQ Analysis/Synthesis

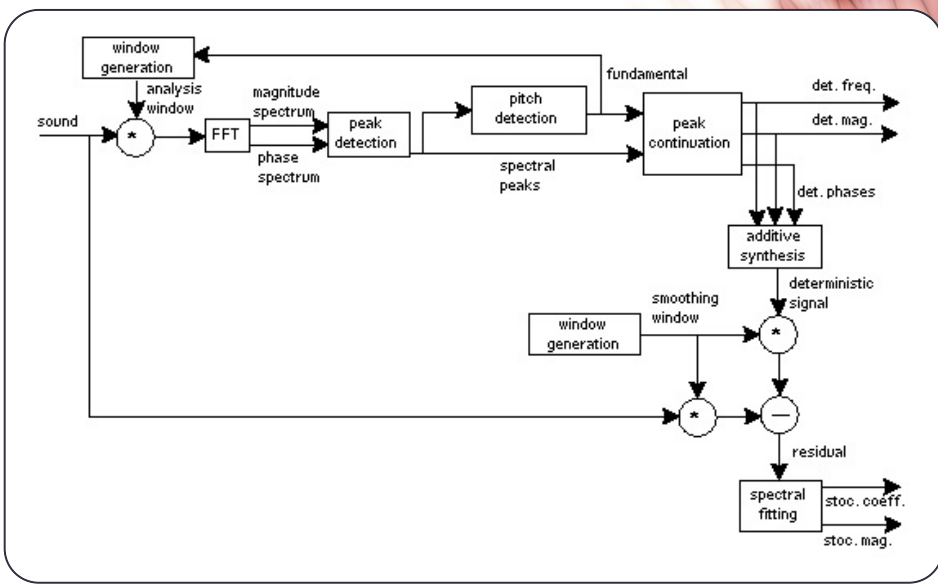
- Because of the spectral representation, time stretching and frequency shifting are easy
- Sinusoid track representation allows phase continuity – no "buzz" at frame rate
- Noise is handled by using large numbers of sinusoids – not very efficient or flexible
- Resonances not modeled
- No obvious cross-synthesis method (interpolation?)

ICM Week 9 25

SMS – Spectral Modeling Synthesis

- Xavier Serra's thesis work, continued at Music Technology Group, Audiovisual Institute, Pompeu Fabra University, Barcelona
- Extends MQ Analysis/Synthesis with explicit model for noise

ICM Week 9 26



SMS Examples

- Original
- Deterministic
- Stochastic
- Synthesis
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11

