


INTRODUCTION TO COMPUTER MUSIC
SPECTRAL INTERPOLATION SYNTHESIS

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

ICM Week 12 Copyright © 2002-2013 by Roger B. Dannenberg 1



**SPECTRAL
INTERPOLATION
SYNTHESIS**

Roger Dannenberg
Marie-Helen Serra
Dean Rubine
Istvan Derenyi
Hank Pelerin
Ning Hu

ICM Week 12 Copyright © 2002-2013 by Roger B. Dannenberg 2

Additive Synthesis and Table Lookup Synthesis

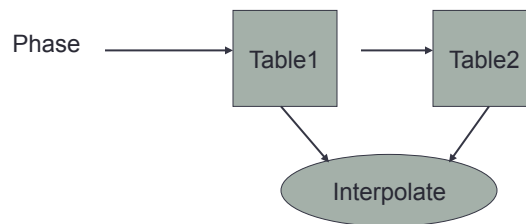
- Additive Synthesis:
 - Every partial has independent frequency and amplitude
 - Analysis/synthesis possible, but no simple parametric control
- Table-Lookup Synthesis
 - Relative amplitudes of all partials are locked in
 - Frequencies are all harmonic

Table Lookup Oscillator Review

```
float table[512] = { ... some waveform ... };
double phase = 0.0;
void osc(double hz, float table[], float out[]) {
    double incr = hz * 512 / sample_rate;
    for (int i = 0; i < block_size; i++) {
        int iphase = floor(phase);
        double x1 = table[iphase];
        out[i] = x1 + (phase - iphase) *
                (table[iphase+1] - x1);
        phase += incr;
        if (phase > 512) phase = phase - 512;
    }
}
```

Spectral Interpolation

- Interpolate between tables
- Keep phases coherent so that interpolation is truly interpolation of spectra
- Restricted to harmonic spectra



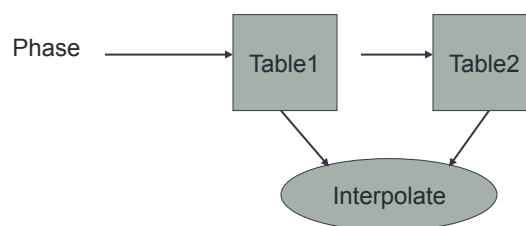
ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

5

Spectral Variation by Interpolation

- Reload tables with new spectra
- Relatively slow update (~20 tables/second)
- Where do tables come from?

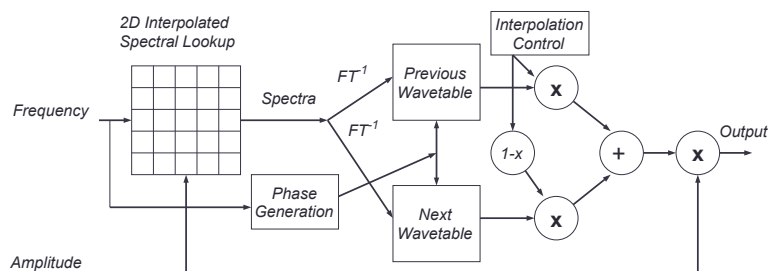


ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

6

Use Pitch and Amplitude to Compute Spectra



ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

7

Measuring Spectra

- We use phase-vocoder style analysis on
- Crescendo at different pitches
- Yields 2-D (pitch, amplitude) tables of spectra

ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

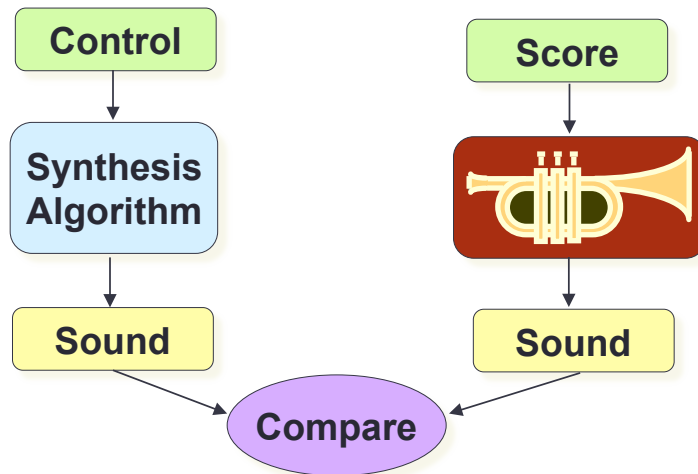
8

Attacks Are Too Rapid And Inharmonic

- Solution: use sampled attacks (~30ms)
- New Problem: How do you join attack to synthesized sound?
 - Cross-fade does not work – too many phase problems
 - Make attack long enough to settle to harmonics
 - Analyze phase of every partial at end of attack
 - Synthesize tables with matching phase
 - Splice with very short cross-fade (2ms) or none

WHERE DOES CONTROL INFORMATION COME FROM?

Traditional Synthesis Research

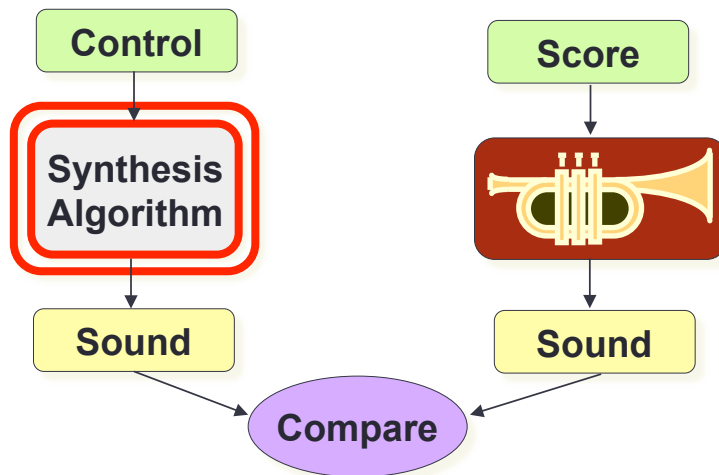


ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

11

Traditional Synthesis Research

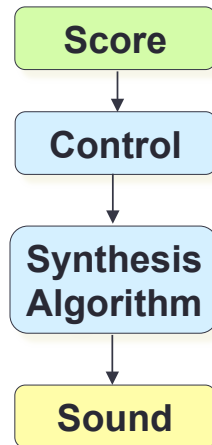


ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

12

SIS Research Approach

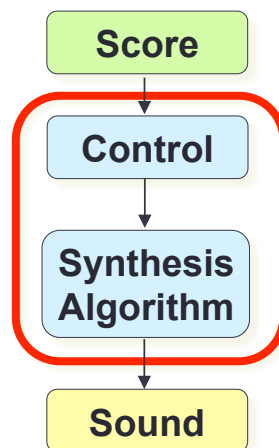


ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

13

SIS Research Approach

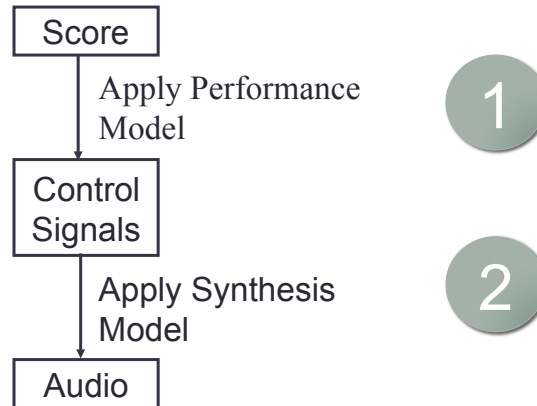


ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

14

Divide-and-Conquer

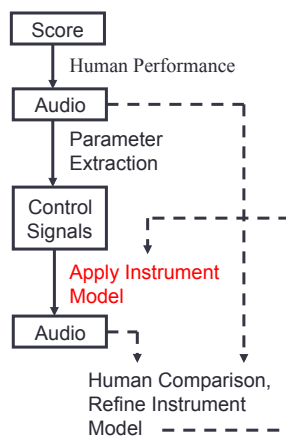


ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

15

Research Model: Synthesis Refinement

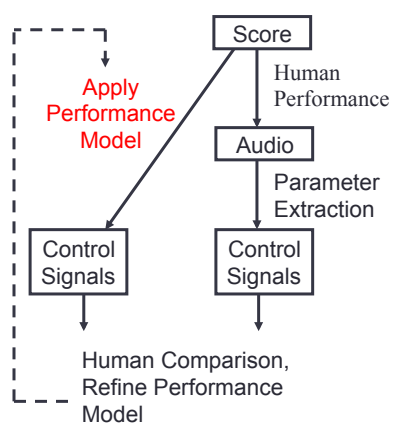


ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

16

Research Model: Control Refinement



A STUDY OF TRUMPET ENVELOPES

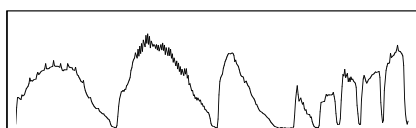
Introduction

- Envelopes:
 - are crucial to sound synthesis;
 - depend upon context.
- We claim:
 - envelopes largely determined by context;
 - envelope generation techniques can improve synthesis.

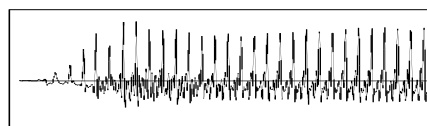
The Big Picture



Score



Control



Sound

Related Work

- Moorer, Grey, and Strawn '78
- Clynes '85
- Chafe '89
- Canazza, De Poli, Roda', and Vidolin '97
- Sundberg, Askenfelt, and Fryden '83
- *Key: consider context, study envelopes in musical phrases*

An Experiment

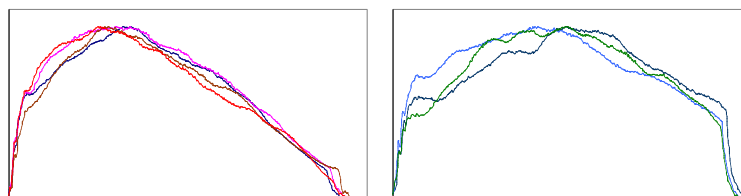
The image displays seven musical phrases on a single staff, each with a circled note. The phrases are arranged in two rows. The first row contains four phrases with dynamic markings *mf*, *mf*, *f*, and *mf*. The second row contains three phrases with dynamic markings *mf*, *p*, and *mf*. The circled notes are: the second note of the first phrase, the first note of the second phrase, the second note of the third phrase, the first note of the fourth phrase, the first note of the fifth phrase, the second note of the sixth phrase, and the second note of the seventh phrase.

Question: How do factors affect the center of mass* (first moment)?

***Why Center of Mass?:** simplest measure of shape.

What Did We Measure?

- Notes and envelopes defined by inter-onset times.
 - because end-of-note not well defined
 - “micro pauses” incorporated into data



ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

23

Some Results

- Up-up phrases showed later center of mass than other combinations
- Large intervals had earlier center of mass than small intervals
- Legato articulation gave later center of mass than others (this should be obvious).

ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

24

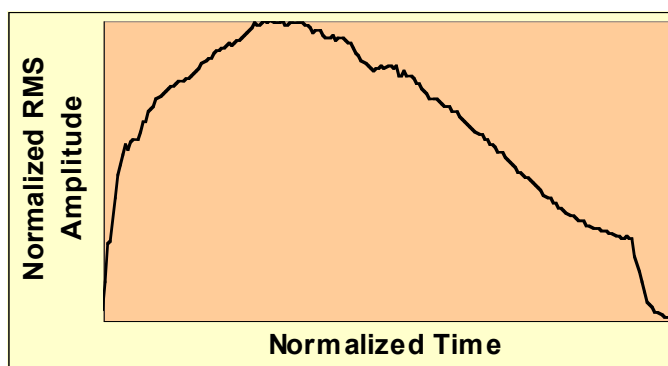
BACK TO SPECTRAL INTERPOLATION SYNTHESIS

ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

25

Looking at “Real” Envelopes

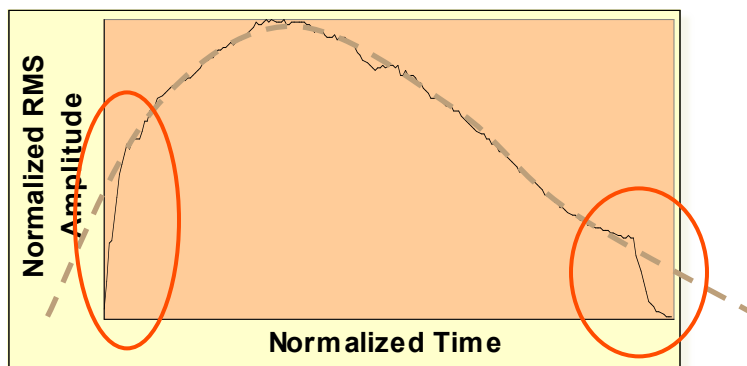


ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

26

Tongue and Breath

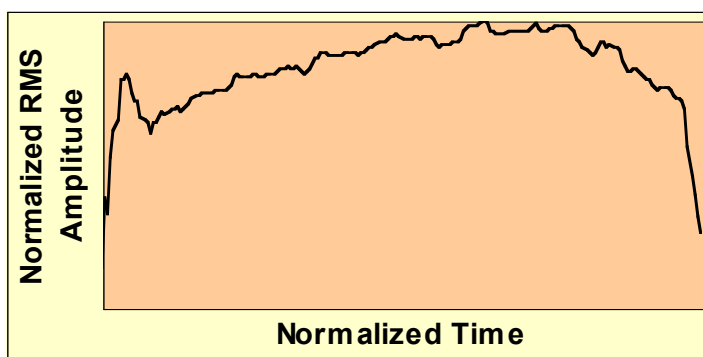


ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

27

Envelope of a Slurred Note

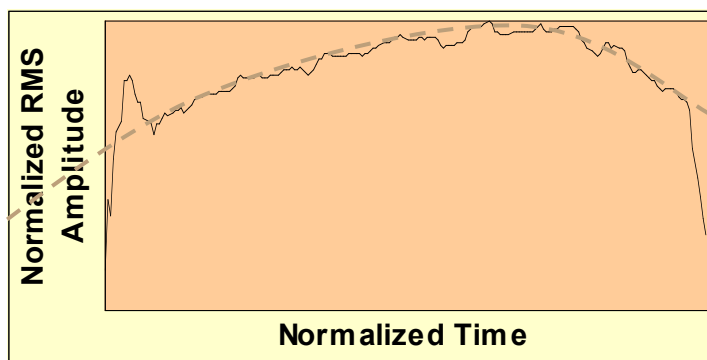


ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

28

Envelope of a Slurred Note



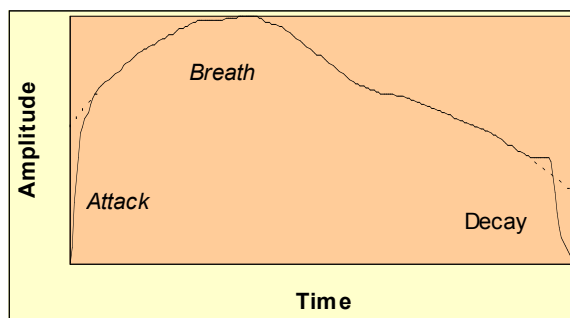
ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

29

Toward an Envelope Model

- Breath envelope gives overall shape
- Tongue articulates beginning and ending



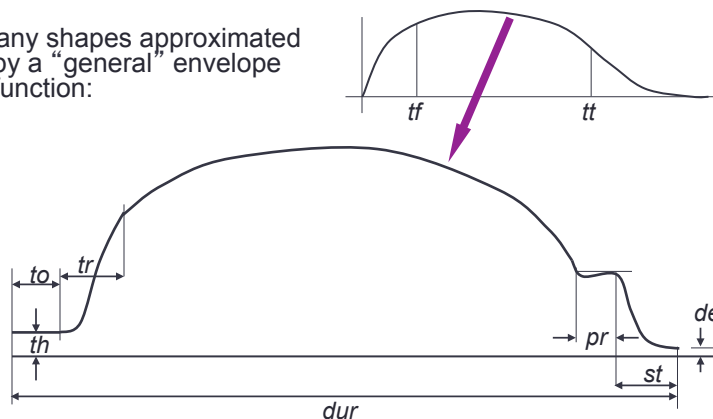
ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

30

Amplitude Envelope Specification

Many shapes approximated by a “general” envelope function:



ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

31

Computing Parameters

- Previous slide shows 9 detailed parameters
- Compute them from:
 - *pitch* (in semitones, according to score)
 - *dur* (in seconds, according to score)
 - *begin-phrase* (is this the first note in a phrase?)
 - *end-phrase* (is this the last note in a phrase?)
 - *from-slur* (is there a slur from preceding note?)
 - *to-slur* (is there a slur to the next note?)
 - *direction-up* (is this note higher than preceding note?)

ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

32

Method

- Score provides “actual” parameters
- Study real performances
- Performance provides envelopes
- Find detailed parameters manually
- Generalize from observed trends

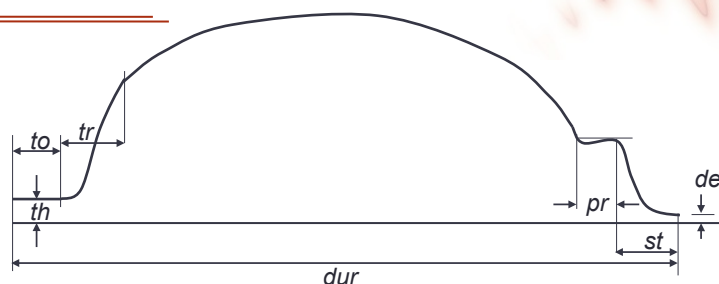
Example: *tf*



```
(cond (from-slur
      (setf takefrom (if direction-up 0.1 0.04)))
      (t
       (setf takefrom (+ 0.03 (* 0.01 (- log-dur))))))
```

		<i>direction-up</i>	
		<i>t</i>	<i>f</i>
<i>from-slur</i>	<i>t</i>	0.1	0.04
	<i>f</i>	0.03 - 0.01 × log ₂ (<i>dur</i>)	

Example: (*st*, *pr*)



end-phrase $\rightarrow (0.02, 0.02)$

to-slur & direction-up $\rightarrow (0.03, dur \times 0.2)$

to-slur $\rightarrow (0.02, 0.001)$

o.w. $\rightarrow (0.06 + 0.005 \times \log_2(dur),$
 $0.03 - 0.01 \times \log_2(dur))$

Envelope from Score

- We use a 9-parameter envelope model
- One “reference” breath envelope is cropped and stretched (2 parameters + duration)
- Breath envelope is multiplied by attack and decay envelopes to derive final envelope
- Envelope depends on: pitch, pitch context, articulation, next note articulation, duration.
- Score-to-envelope mapping is hand-crafted.

Frequency Envelope

- Use analyzed frequency envelopes
- Capture natural frequency variation
- Pitch dependent
- No vibrato (yet)

Sound Examples

- CSIS Synthesis Example
- Another CSIS Synthesis Example
- CSIS (Without Accompaniment)



Summary

- Modeled from “real” performance
- Phrase-at-a-time
- Melodic direction (lookahead)
- Articulation (lookahead)
- Duration
- Pitch
- Currently rule-based, *ad-hoc*

Conclusions

- Envelopes are critical to music synthesis
- Statistically valid relationships between score parameters and envelope shape
- Breath + Tongue Model
- *Study of musical phrases and notes in context is critical to future synthesis research.*

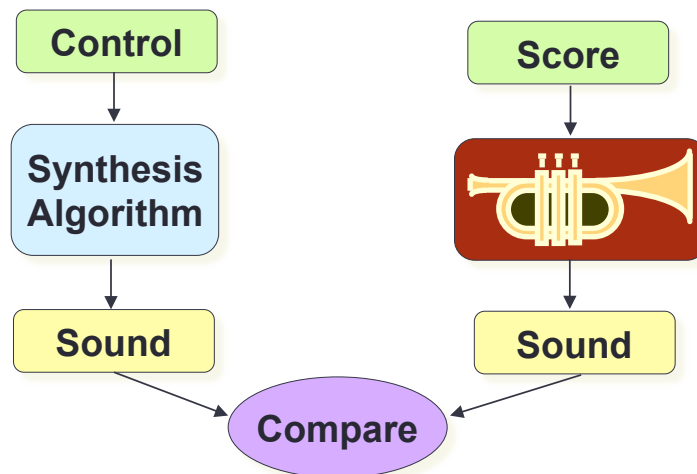
Implementation

- SNDAN (Beauchamp et al.) for spectral, RMS, and frequency analysis
- Nyquist (www.cs.cmu.edu/~rbd)
 - score representation
 - compute envelope parameters
 - cut, splice, shape envelopes
 - amplitude, frequency curves → spectral interpolation synthesis

So What?

- Based on these examples (and others), I claim we have an adequate “Synthesis Algorithm.”
- What happens if we return to the “Traditional Synthesis Research” model with the CSIS synthesis algorithm?

Traditional Synthesis Research





ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

43

SIS With “Standard” Envelopes

- Amplitude, Time, Duration are correct
- Envelope shape is “correct”
- Timbre is correct
- Two Envelopes (better slurs) 
- SIS model with computed envelopes 

ICM Week 12

Copyright © 2002-2013 by Roger B. Dannenberg

44

Discussion

- *Since the only “wrong” ingredient is envelope shape, envelope shape must be critical for synthesis.*
- This observation could only be made after we perfected (?) trumpet synthesis.

More Discussion

- It follows from our results that...
Previous work on synthesis was wrongly assuming simple envelope templates are sufficient.
- In other words...
If you had a good synthesis algorithm, how would you know it?

Still More Discussion

- Equal time for the physical modelers:

My goal is to develop sufficiently rich instruments that musicians will want to learn how to control them.

(paraphrasing Julius Smith)