A Continuous Music Keyboard
Controlling Polyphonic Morphing
Using Bandwidth-Enhanced Oscillators

Lippold Haken, Kelly Fitz, Ed Tellman, Patrick Wolfe, Paul Christensen

CERL Sound Group, University of Illinois, http://datura.cerl.uiuc.edu

Abstract

The Continuum is a new type of polyphonic music performance device that is approximately the same size as a traditional synthesizer keyboard, but has a continuous playing surface instead of discrete keys. It tracks independent x, y, and z (pressure) position for up to 10 simultaneous notes. The x, y, z position from each finger is used to control bandwidth-enhanced oscillators in a polyphonic timbre morph. Real-time synthesis is based on non-real-time timbral analyses done with Lemur.

1 Introduction

The Continuum employs new keyboard hardware, new analysis techniques, and new synthesis algorithms to give a performer control over a real-time polyphonic timbre morph.

The Continuum uses the x position of each finger as a continuous pitch control for a note. One inch in the x direction corresponds to a pitch change of 200 cents. The performer must place fingers accurately to play in tune, and can slide or rock fingers for pitch glide and vibrato.

The y position of each finger provides timbral control for each note in a chord. The performer can bring out certain notes in a chord by playing them with a different timbre; this requires careful y placement of the fingers. By sliding fingers in the y direction while notes are sounding, the performer can create timbral glides.

The z (pressure) of each finger provides dynamic control. Changing finger pressure is used to create tremolo. An experienced performer may simultaneously play a crescendo and decrescendo on different notes.

The Continuum’s sound synthesis algorithms make complex spectral changes in response to the performer’s finger movements. The x, y, and z position of each finger corresponds to a location in a three-dimensional “timbre space.” The position controls the timbral mix, or morph, between previously analyzed (or previously synthetically generated) sounds in the timbre space. Timbre morphing is the process of combining several sounds to create a new sound with intermediate timbre. This process differs from simply mixing sounds, as only a single sound, with some of the characteristics of each of the original sounds, is audible as the morphed sound. By changing finger position, the performer changes the contribution of timbres to the morph. If the performer creates vibrato by rocking a finger in the x direction, this causes a change in pitch as well as the appropriate spectral changes in each vibrato cycle.

The Continuum sound synthesis utilizes bandwidth-enhanced oscillators [1] which are capable of spreading energy around each sinusoidal component (spectral line widening [2]). The timbres are pre-analyzed using Lemur, and the real-time synthesis is done in Symbolic Sound’s Kyma environment [3,4].

2 Previous Devices

Modern electronic keyboards allow the performer to use key velocity and polyphonic aftertouch to control sound synthesis. These capabilities are extended by certain experimental keyboards, such as Moog’s clavier [5]. Moog’s clavier measures not only pressure aftertouch, but also other parameters including the exact horizontal and vertical location of each finger on the surface of each key.

Researchers at McGill University demonstrated the Rolky at the International Computer Music Conference in 1985 [6]. It had a continuous performance surface, and tracked each finger's position. The Rolky gave the performer extensive control over each note in the chord. In the early 1980’s we began working on the Continuum. Like the Rolky, the Continuum has a continuous performance surface rather than discrete keys. Our early attempts at the Continuum were of limited success, due to mechanical problems as well as limitations in our synthesis algorithms.
3 Continuum Playing Surface

*Figure 1* shows the mechanical design of the Continuum playing surface [7]. Hall-effect sensors measure the position of the magnets. When the performer applies finger pressure, the rods under the finger are depressed, and the magnets on those rods move closer to the sensors.

Figure 1: The Continuum uses Hall-effect sensors to detect the positions of magnets mounted on each rod.

Scanning software running on a Macintosh computer detects fingers by looking for any bar which has normalized pressure values greater than both of its neighboring bars. We call this the center bar, and the neighboring bars the right bar and left bar. The pitch (x position) is computed from the position of the center bar. The front-to-back position (y position) is computed by summing the back sensors on the left, center, and right bars, and dividing by the sum of all six sensors on these bars. The total finger pressure (z position) is the sum of all six sensors.

The Continuum tracks the exact pitch (x position) of a finger using parabolic interpolation. A parabola is drawn through the left, center, and right bars’ normalized sensor values. *Figure 2* shows this, as well as the detection of a vibrato. In this example the center bar is always the same bar – it is always more depressed than its right and left neighbor. Still, the x position is accurately tracked as the performer rocks the finger back and forth, because the movement of the neighboring bars affects the minimum point of the parabola.

4 Timbre Space

The x, y, and z position for each finger determine a location in a three dimensional timbre space. We divide the three dimensional timbre space into cubes; neighboring cubes share one face [8]. *Figure 3* shows one cube of a timbre space which was made from Lemur analyses of four cello tones and four trombone tones. Each corner of the cube has a set of amplitude, frequency, and bandwidth envelopes derived from an analyzed tone. These 8 sets of envelopes completely define this part of the timbre space.

Figure 3: One cube in a timbre space created from Lemur analyses of cello and trombone tones.

The timbre space we will demonstrate at ICMC97 is made of 24 cubes based on the Lemur analyses of 78 tones (39 trombone tones and 39 cello tones). When the performer places a finger on the Continuum, the finger's x, y, and z location falls within one of the cubes in the timbre space. If the location of a finger gradually changes during a note, this corresponds to a gradual change of location within a cube of the timbre space. If the x, y, and z location changes greatly during a note, the timbre space location of the sound may travel through the face of one cube into a neighboring cube.

Figure 2: Exact pitch tracking using parabolic interpolation: A finger rocking left, centered, and rocking right.
5 Lemur Analysis

We pre-analyze tones using Lemur [1]. Lemur uses enhancements to the McAulay-Quatieri (MQ) sinusoidal representation [9] to better accommodate “noisy” signals. Most instrument tones contain noise in their attacks and transients; this noise is not spread evenly over the spectrum, but varies with frequency. Purely sinusoidal models (such as MQ) represent the noise with an abundance of components whose amplitudes and frequencies fluctuate wildly, yielding a representation which is difficult to manipulate.

The MQ representation consists of amplitude and frequency envelopes for each partial, computed using parabolic interpolation of the magnitude spectrum around spectral peaks [10]. Lemur uses a modified spectral amplitude which is the same as the MQ amplitude for strongly sinusoidal signals, and somewhat greater than the amplitude of the spectral peak for noisy signals. Lemur additionally specifies a bandwidth, or “noisiness” envelope for each partial.

Lemur divides the short-time frequency spectrum into regions in order to associate noise energy with nearby sinusoidal components. Each region contains a single strong magnitude peak, selected according to the usual MQ process. All of the spectral energy in a region is associated with a single, bandwidth-enhanced component. The frequency of this component is found using parabolic interpolation on the magnitude spectrum. Its amplitude is equivalent to that of a sinusoid containing all the region’s spectral energy, and its bandwidth factor specifies the ratio of noise energy to sinusoidal energy.

The Lemur method of energy association is not analytically rigorous, in the sense that the noise energy associated with each partial is only an approximation of the spectral energy in the partial’s frequency region. However, the approximation is reasonable for signals having mostly sinusoidal energy.

Existing techniques, such as the stochastic modeling of Serra and Smith [11], represent the noise energy and sinusoidal energy of signals as distinct model components. Lemur retains the homogeneity of the purely sinusoidal representation by extending it with bandwidth envelopes. This homogeneity is advantageous for the timbral manipulations used in the Continuum’s polyphonic morphing.

6 Real-time Synthesis

As fingers move on the Continuum, their positions are tracked by the Continuum sensors and software on the Macintosh. From each finger’s x, y, and z information, the Macintosh selects which cube in the timbre space contains each finger’s location. The rest of the real-time computations are implemented on the Capybara signal processor with the Kyma graph shown in Figure 4.

Figure 4: Kyma graph for real-time synthesis with the Continuum. Only the subgraph for the sixth voice is shown.
For each finger, the Capybara computes the new envelope time index, and does any time stretching or compression required for the timbres being morphed [12]. The Capybara synthesizes between 20 and 80 partials for each finger, depending on the timbre. For each partial, 24 envelopes pre-analyzed by Lemur (8 amplitude, 8 frequency, 8 bandwidth) are read by eight AmpFreqBandwidthEnvs Sound Objects. These Sound Objects interpolate between the pre-analyzed envelope values, to avoid problems under extreme time stretch conditions, and output parameter streams [13].

The streams are weighted according to the x, y, z location of the finger within the cube, and summed together to make a single morphed stream. This morphed stream is then combined with the morphed streams for other fingers using a Concentrator Sound Object. The Concentrator Sound Object will perform partial pruning [10] if many fingers are simultaneously pressing on the Continuum. The Concentrator outputs a single envelope parameter stream which has the morphed partials for all the fingers; this stream feeds into cascaded BWEOscillatorBank Sound Objects. Each BWEOscillatorBank implements bandwidth-enhanced oscillators for a subset of the partials. The BWEOscillatorBanks also do envelope exponentiation and slope generation, so that amplitude, frequency, and bandwidth values are updated at every sample time. The outputs of these cascaded BWEOscillatorBank objects are summed; this is the final audio output.

It is interesting to note that the computation of the bandwidth-enhanced sinusoids takes a minority of the processing power in synthesis. The envelope processing and timbre morphing is the majority of the work, since it involves processing 15,360 envelopes in real time (8 amplitude envelopes + 8 frequency envelopes + 8 bandwidth envelopes = 24 envelopes to compute and morph for each partial; 64 partials per timbre; and 10 simultaneous notes for 10 fingers: 24 * 64 * 10 = 15,360).

References


