SCANNED SYNTHESIS

Bill Verplank, CCRMA--Stanford University
Max Mathews, CCRMA--Stanford University
Robert Shaw, Santa Fe Institute
contact: <verplank@ccrma.stanford.edu>

ABSTRACT

This paper describes a new technique for the synthesis of musical sounds which we have named Scanned Synthesis. Scanned Synthesis is based on the psychoacoustics of how we hear and appreciate timbres and on our motor control (haptic) abilities to manipulate timbres during live performance. A unique feature of scanned synthesis is its emphasis on the performer's control of timbre. Scanned synthesis involves a slow dynamic system whose frequencies of vibration are below about 15 hz. The system is directly manipulated by motions of the performer. The vibrations of the system are a function of the initial conditions, the forces applied by the performer, and the dynamics of the system. Examples include slowly vibrating strings and two dimensional diffusion equations. To make audible frequencies, the "shape" of the dynamic system, along a closed path, is scanned periodically. The pitch is determined by the speed of the scanning function. Pitch control is completely separate from the dynamic system control. Thus timbre and pitch are independent. This system can be looked upon as a dynamic wave table controlled by the performer.

1. PSYCHOPHYSICAL BASIS

The psychophysical basis for Scanned Synthesis comes from our knowledge about human auditory perception and human motor control abilities. In the 1960's Risset showed that the spectra of interesting timbres must change with time. We observe that musically interesting change rates are less than about 15 hz which is also the rate humans can move their bodies. We have named these rates Haptic rates.

In the middle 1960's Jean-Claude Risset (1969a, 1969b) demonstrated that in order to make good simulations of traditional instruments the spectrum must change with time over the course of a note. For example, in a brass timbre, the proportion of high frequency energy in the spectrum must increase as the intensity of the sound increases at the beginning (attack part) of a note.

2. HAPTIC FREQUENCIES

Over the last decades, many extensions of Risset's work led to a better understanding of the properties of spectral time variations that the ear hears and the brain likes. Spectral time variations can also be usefully characterized by their frequency spectrum. These frequencies are much lower (typically 0 to about 15hz) than audio frequencies (50hz to 10000hz).

Either by a happy accident of nature or because of the way human beings are built, the frequency range of spectral changes the ear can understand is the same as the frequency range of movements of our body parts--arms, fingers, articulators, etc--that we can consciously control. Scanned synthesis provides methods for directly manipulating the spectrum of a sound by human movements.

At present the terminology with which to describe spectral time variations is not well established. Some kinds of spectral time variations, particularly vibrato and tremolo, are called modulations. But other kinds, such as occur in brass timbres are unnamed. We here propose the name *haptic frequencies* to characterize these variations.

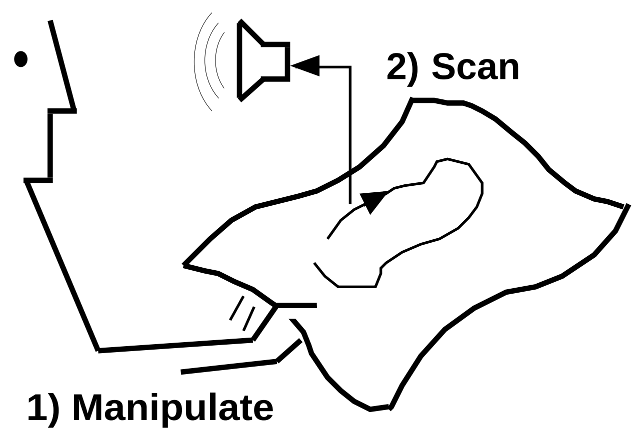

Figure 1. *Scanned Synthesis consists of 1) manipulating a dynamic system and 2) scanning out a wave-shape from along a path.*

3. SCANNED SYNTHESIS

The essence of scanned synthesis is to use a slowly vibrating object whose resonant frequencies are low enough so the performer can directly manipulate the object's vibrations by motions of his body and to scan (measure) the shape of the object along a periodic path by a periodic scanning function whose period is the fundamental frequency of the sound we wish to create. The scanning function translates the slowly changing spatial wave shape of the object into a sound wave with audio frequencies which the ear can hear.

Scanned synthesis can be looked upon as a descendent of wave table synthesis. In wave table synthesis, points in a function of one independent variable are computed and stored in successive memory locations in a computer. This chunk of memory (the wave table) is scanned or read by a periodic scanning function to produce the samples of the audio sound wave. The period of the scanning function is the period of the synthesized sound. The scanning process is computationally simple and efficient. The computation of the wave table need only be done once, and thus can be computationally intensive.

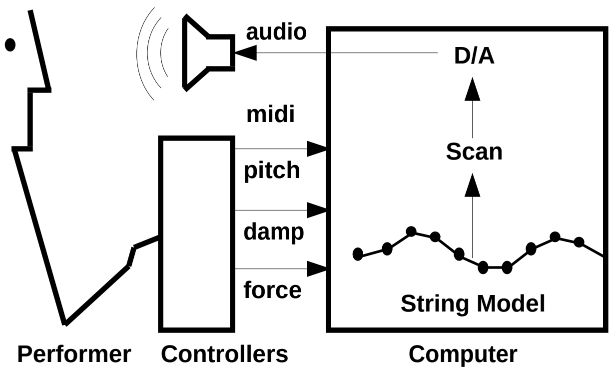

Figure 2 *A performer uses a variety of controllers to send pitch to the scan rate and parameters (e.g. damping) and disturbances (e.g. force) to the model (generator).*

A general block diagram of our model is shown in Fig 2. It is a real-time program which generates a sequence of samples that are read out of the computer at the audio sampling rate (typically 44100 samples per second) through a d-to-a converter and sent to a loud speaker. It also contains input channels through which the performer "plays" the model. The input channels (typically midi) are connected to physical controllers that the performer touches and manipulates. These can include midi keyboards, radio-batons, and Phantom sticks. We have also used a video camera as an input device.

The haptic generator in the model generates sampled spatial frequencies. The scanning path, becomes an array of numbers in the computer memory. This array can be looked upon as a dynamic wave table. These numbers are changed at haptic rates by the haptic generator. Thus the numbers are functions of both time and position in the wave table. These numbers are scanned, ie read out along their position in the wave table, by a periodic scanning function whose period is an audio frequency (for example 1/440 second). The resulting samples are sent to the d-to-a converter.

Although the scanning path is a 1-dimensional path, the haptic model itself can have more than 1-dimension. For example inputs from a video camera are processed by a 2-dimensional model.

In order to be useful, the numbers in the dynamic wave table must represent useful spatial frequencies. It is not sufficient that a number in given location in the array change in time at haptic frequency. An individual number must be related to its neighbors along the scanning path in a way which represents a desired spatial frequency. This spatial frequency is converted to a time frequency by the scanning function. This property is achieved by the choice of mathematical functions computed by the haptic generator.

4. ONE-DIMENSIONAL STRING MODEL

A useful model which generates spatial frequencies can be derived from the finite element approximation to a string. It can be thought of as a string of masses connected by springs. The equations of motions of the masses can be simply derived from Newton's equations. The resulting differential equations can be approximated by difference equations which can be solved by a computer as shown in Appendix A.

We have studied Scanned Synthesis chiefly with a finite element model of a generalized string. Our finite element models are a collection of masses connected by springs and dampers. We have generalized a traditional string by adding to each mass a damper and a spring connected to earth. All parameters -- mass, damping, earth spring strength and string tension -- can vary along the string. The performer manipulates the model by pushing or hitting different masses and by manipulating parameters.

Musical applications of finite element models were pioneered in the 1970's by Cadoz (1978, 1979, 1993) and his associates. Our work differs from that of Cadoz in that our models vibrate at low (haptic) frequencies and must be scanned to obtain audio frequencies. The models of Cadoz generally vibrated directly at audio frequencies. This difference is important. The performer can directly manipulate the motions of our haptic models. Also, slow models also require much computer power.

We have generalized the string model by:

1) allowing each element to have an arbitrary mass, M_i,

2) attaching damping, D_i from each mass to "earth",

3) attaching "centering spring" C_i from each mass to "earth",

4) and applying a haptic force, f_i to each mass.

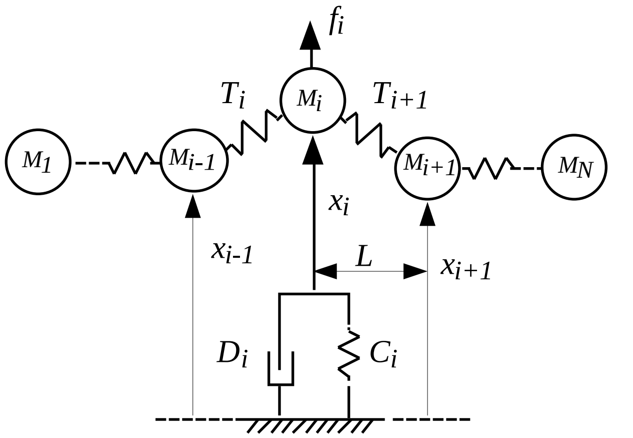

Figure 3 *Finite Element Model of string: M, mass; T, spring stiffness between masses; C, spring to earth; D, damping to earth; L, length between masses; x, position; f, force. For a circular string,* M_N *connects to* M_1*.*

5. EXPERIMENTAL RESULTS

We have tried a number of different ways of playing the string model. These include giving the string an initial displacement and releasing it (plucking), giving the string an initial velocity but no displacement, and applying haptic forces controlled by the performers hands to various masses of the string model. All these techniques can produce interesting timbres. In general the most interesting sounds involve nonuniform strings in which the damping, the tension, and the centering springs vary along the string. In most conditions the performer also can vary several of these parameters in time during the course of a note.

6. SPECTRAL MODULATION DEMONSTRATION

Since the scanning process is independent of the computation of the string shape, it is possible to stop modifying the string shape (freeze the string) but continue scanning and hearing the now unvarying spectrum. The effect is dramatic. Within seconds after the string is frozen, the timbre becomes dull and uninteresting. We interpret this demonstration as supporting our basic hypothesis that a spectrum changing at a haptic rate is essential to an interesting timbre.

7. TWO-DIMENSIONAL MODELS.

We have produced interesting timbres with various two dimensional objects and various equations including a gong, a set of coupled strings and the heat equation with boiling. We have also used chaotic equations such as the Kuramoto-Shivashinski equation. Space does not allow describing this work.

8. CONCLUSIONS

The fundamental elements in Scanned Synthesis are:

1) A dynamic system which has slow modes of vibration.

2) Controllers allowing a performer to manipulate the system at haptic rates.

3) A scanning process which periodically scans the dynamic system at audio rates.

Our initial demonstrations have produced interesting timbers that sound live, as do traditional instruments, but do not resemble traditional instrument timbres. Our initial work has only barely started to explore obvious possibilities in Scanned Synthesis models.

We believe that scanned synthesis may provide a way of "performing" timbre so it will be a major structural element of future music such as is pitch in present music.

REFERENCES

Boulanger, R. 2000a “Scanned Synthesis & CSound @ CSounds.com”, <http://www.csounds.com/scanned>.

Boulanger, R., Smaragdis, P. 2000b. “Scanned Synthesis: An Introduction and Demonstration of a New Synthesis and Signal Processing Technique” , ICMC, Berlin.

Cadoz, C & Florens, J.L. 1978 "Fondements d'une demarche de recherche informatique/musique", *Revue d'Acoustique* No45, pp. 86-101. Paris.

Cadoz, C. 1979 "Synthese sonore par simulation de mecanismes vibratoires", These de Docteur Ingenieur, Specialite Electronique - I.N.P.G. Grenoble.

Cadoz, C, Luciani, & Florens, J.L. 1993 "Cordis-Anima: a Modeling and Simulation System for Sound and Image Synthesis - The General Formalism", *Computer Music Journal*, Vol 17-1, MIT Press.

Massie, Thomas H. and J. K. Salisbury. 1994 "The PHANTOM Haptic Interface: A Device for Probing Virtual Objects." Proceedings of the ASME Winter Annual Meeting, Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, Chicago, IL.

Risset, J.-C, and M. Mathews. 1969a. “Analysis of Instrumental Tones.” *Physics Today* 22(2): 23-30.

Risset, J.-C. 1969b “An Introductory Catalog of Computer-synthesized Sounds.” Murray Hill, New Jersey: Bell Laboratories.

APPENDIX A
Equations for a circular string.

This appendix derives a linear constant coefficient difference equation, starting with Newton's equations of motion for a generalized string. For our purposes it is not so important that the approximate solution be accurate than that it be stable and musically interesting. Our simulations have shown these difference equations to be successful for these purposes.

Newton's equations can be written for each element (i) as a set of integral equation for acceleration (a_i), velocity (v_i) and position (x_i).

$$a_i = \frac{1}{M_i}\left[K_i\left(x_{i-1} - 2x_i + x_{i+1}\right) - C_i x_i - D_i v_i + f_i\right]$$

$$v_i = \int a_i dt$$

$$x_i = \int v_i dt$$

Functions of time:

$a_i = a_i(t)$, acceleration of the i th element,
$v_i = v_i(t)$, velocity of the i th element,
$x_i = x_i(t)$, position of the i th element,
$f_i = f_i(t)$, (haptic) force on the i th element.

Model parameters:

M_i mass of i th element,
$K_i = \frac{T_i}{L_i}$ effective spring constant between i and i-1.
C_i spring constant to earth for the i th element,
D_i damping of the i th element.

Boundary conditions for a string with two fixed ends:

$$x_0 = v_0 = 0$$
$$x_N = v_N = 0$$

For a circular string with no end (not physically realizable):

$$a_1 = \frac{1}{M_1}\left[K_1\left(x_N - 2x_1 + x_2\right) - C_1 x_1 - D_1 v_1 + f_1\right]$$

$$a_N = \frac{1}{M_i}\left[K_N\left(x_{N-1} - 2x_N + x_1\right) - C_N - D_N + f_N\right]$$

We show the numerical solutions in the form of computer code. Time sampling equation: j*S = t_j

x0[i]=$x_i(t)$, x1[i]=$x_i(t-S)$, x2[i]=$x_i(t-2S)$
v0[i]=$v_i(t)$, v1[i]=$v_i(t-S)$, v2[i]=$v_i(t-2S)$
a0[i]=$a_i(t)$, a1[i]=$a_i(t-S)$, a2[i]=$a_i(t-2S)$

Equation A1 becomes:

```
a0[i]=(K[i]*(x1[i-1]-2*x1[i]+x1[i+1])
      -C[i]*x0[i]-D[i]*v0[i]+h[i])/M[i]
v0[i]=v1[i]+a0[i]*S
x0[i]=x1[i]+v0[i]*S
```

Solving in terms of x only:

```
x0[i]=P1*x1[i]+P2*(x1[i-1]+x1[i+1])
 +P3*x2[i]+P4*h[i]
```
eqA2

where:

```
P1[i]
 =(2+S*D[i]/M[i]2*S*S*K[i]/M[i])/denom
P2[i] = (S*S*K[i]/M[i])/denom
P3[i] = 1/denom
P4[i] = (S*S/M[i])/denom
denom = (1-S*D[i]/M[i]+S*S*C[i]/M[i])
```

Eq A2 can be iterated to give an approximate solution to Newton's equations.

APPENDIX B Scanned Synthesis Methods Explored

This is a brief history of scanned-synthesis, noting the separate approaches to control, excitation, model, and scan-path in rough chronological order. All were carried out at Interval Research over the years 1998-2000.

(timbre) controller	***model***	***excitation***	***scan***	***researcher***
Phantom (Massie, 1994)	string, cochlea	force on one element	forward and back	Verplank
Phantom	plate	force on one element	double spiral	Verplank
Camera	2-D diffusion	intensity on velocity	linear	Shaw
none	boiling	constant heat	linear	Shaw
knob-box	K-S (chaos)	none	linear	Shaw
Radio Baton	string	velocities, positions	circular	Mathews
Keyboard	eight strings	"hammers" on velocity	circular	Verplank
Keyboard	string	samples on position	circular	Cook
"Snake"	strings	velocity	circular	Cook
Keyboard	general CSound	wave forms on position	general	Smaragdis (2000b)