International Symposium on The Creative and Scientific Legacies of Iannis Xenakis,
8-10 June 2006, University of Guelph, Toronto, Canada

The evolution of granular synthesis: an overview of current research

Curtis Roads
Media Arts and Technology
University of California
Santa Barbara CA 93106-6065 USA

9 June 2006

This lecture presents an overview of several projects pursued over the past five years in our laboratories at Santa Barbara. All this research is based on a scientific model of sound initially proposed by Dennis Gabor (1946), and soon afterward extended to music by Iannis Xenakis (1960). Granular analysis (also called *atomic decomposition*) and granular synthesis have evolved over more than five decades from a paper theory and primitive experiments into a broad range of applied techniques. Specific to the granular model is its focus on the microacoustic time scale (typically 1 to 100 ms). Granular methods treat sound as a stream of acoustic particles in both the time domain and the time-frequency (TF) domain. For more details, see Roads (2002; Kling, et al. 2005).

In this lecture, I first very briefly trace the history of the idea of sound particles. Next I will demonstrate PulsarGenerator, an application developed by Alberto de Campo and me in 2001 for a specific type of particle synthesis with links to past analog techniques. I will also demonstrate the SweepingQGranulator, a tool that I wrote in the SuperCollider language for the microfiltration of granulated sound.

The latest threads in this line of research go in two directions. The first is a time-frequency analysis method known as matching pursuit decomposition. The second is a new prototype for generalized synthesis and control of particle synthesis called EmissionControl.

Finally, I would like to demonstrate some of the visualizations that we have been developing in conjunction with this research, some of which are motivated by scientific aims, others of which are artistically motivated, and some that attempt to satisfy both aims.

History of the granular model

The idea that a continuous tone could be decomposed into smaller quantities of time emerged from ancient atomistic philosophies. In the latter part of the fifth century BC, the Greek philosophers Leucippus and Democritus taught that all matter consists of atoms separated by empty space. They speculated that any substance or energy could be divided into smaller and smaller pieces and would eventually reach a point where it could no longer be divided: the atom. Another atomist, Epicurus, founded a philosophical school in Athens and taught his doctrines to a devoted body of followers. Later the Roman Lucretius wrote a poem *De Rerum Natura* that delineated the Epicurean philosophy.

In the seventeenth century, at the dawn of early modern science, the French thinkers Pierre Gassendi and René Descartes revived atomism. A confluence of intellectual energy, emanating from Descartes, Galileo, Beekman, Mersenne, Gassendi, Boyle, and others, gradually forced a paradigm shift away from the Aristotelian philosophical point of view toward a more experimental perspective. Issues in acoustics were central to the growth of science in Western Europe.

The modern concept of sound particles can be traced to Einstein's phonons, which he predicted in 1907. But the phonons consist of inaudible packets of ultrasonic energy at feeble amplitudes. It was Einstein's pupil, Dennis Gabor, who in the 1940s had the fundamental insight that brought the particle model into the domain of perceived sound. The composer Iannis Xenakis learned of Gabor's experiments, and in 1959 he made an experiment in which he approximated granular synthesis by means of tape splicing. He did not, however, continue research in this direction. His book *Formalized Music* described a theory of granular synthesis. It was this description that led to me realize the first implementation of granular synthesis on a computer in 1974.

<Play examples of early granular synthesis from *Microsound*>

Today granular synthesis is a staple of electronic and computer music technique. What are some of the latest developments in this rich vein of research?

Pulsar synthesis: the Pulsar Generator program

Pulsar synthesis (PS) is a powerful method of digital sound synthesis with links to past analog techniques (Roads 2001). PS melds established principles within a new paradigm. In its basic form, it generates electronic pulses and pitched tones similar to those produced by analog instruments such as George Jenny's Ondioline and the Hohner Elektronium, which were designed around the principle of filtered pulse trains. Pioneering electronic music composers such as Karlheinz Stockhausen and Gottfried Michael Koenig used filtered impulse generation as a staple in their studio craft. Pulsar synthesis is a digital technique, however, and so it accrues the advantages of precise programmable control, waveform flexibility, graphical interface, and extensibility. In its advanced form, pulsar synthesis generates a world of rhythmically structured crossbred sampled sounds.

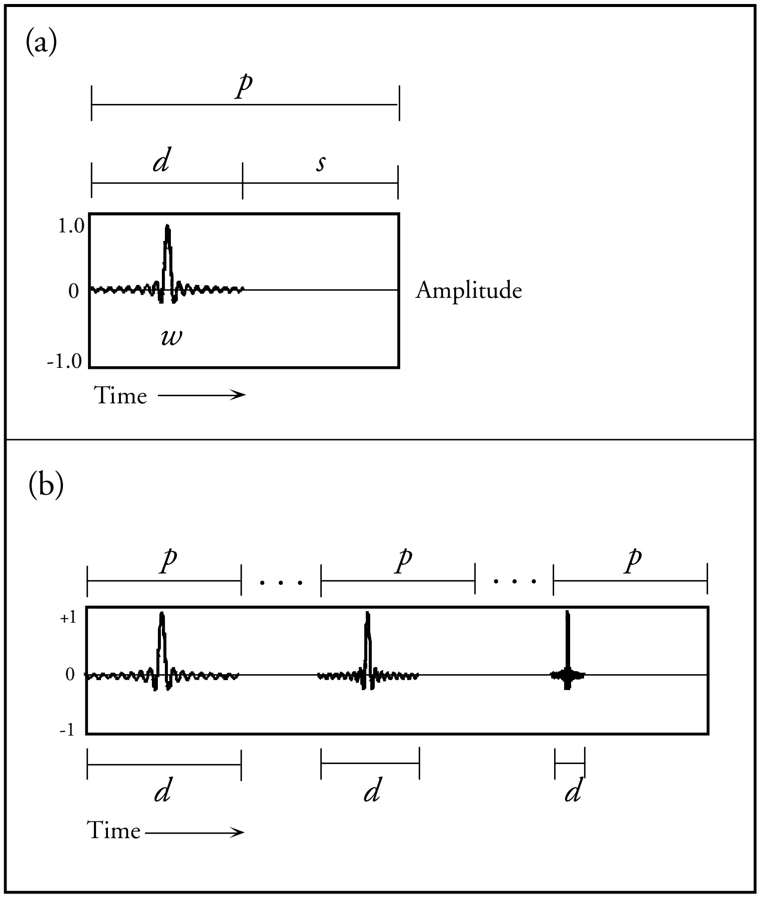


Figure 1. Pulsar synthesis.

In Figure 1(a) we see a single pulsar is a particle of sound. It consists of an arbitrary *pulsaret* waveform w with a period d followed by a silent time interval s. The total duration of a pulsar is $p = d + s$, where p is the *pulsar period*, d is the *duty cycle*, and s is silent. Repetitions of the pulsar signal form a *pulsar train.* Let us define the frequency corresponding to the repetition period as $f_p = 1/p$ and the frequency corresponding to the duty cycle as $f_d = 1/d$. Typical ranges of f_p are between 1 Hz and 5 kHz, and the typical range of f_d is from 80 Hz to 10 kHz.

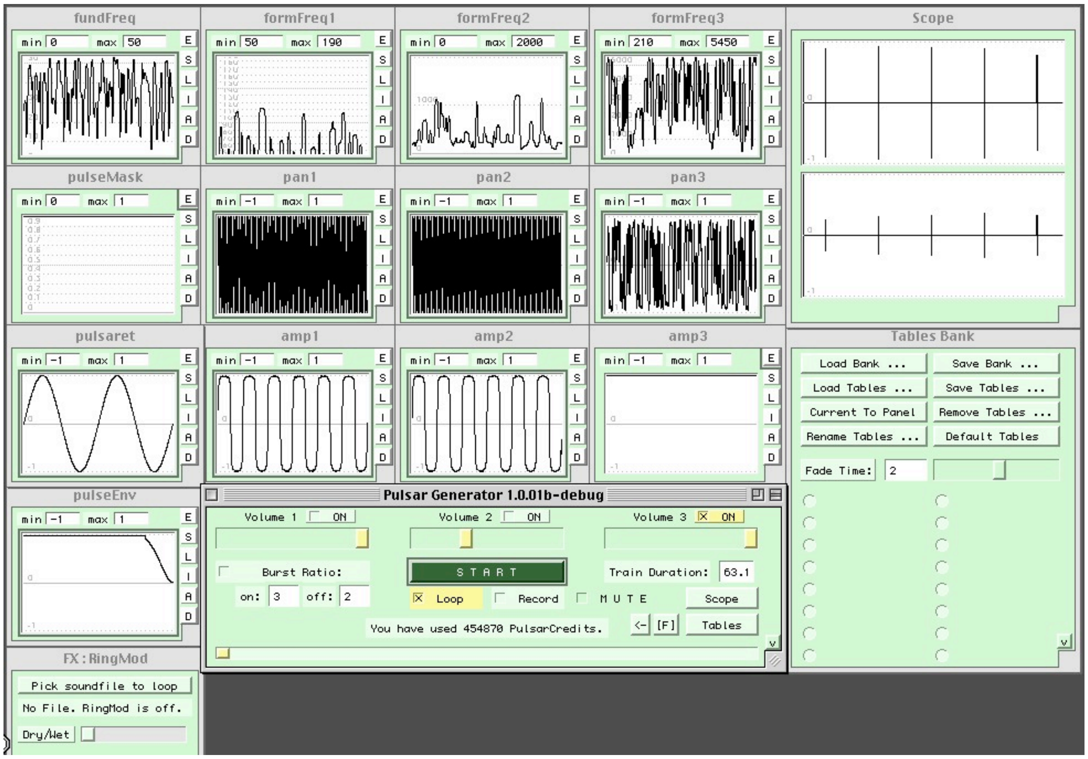
In pulsar synthesis, both f_p and f_d are continuously variable quantities. They are controlled by separate envelope curves that span a train of pulsars. The train is the unit of musical organization on the time scale of notes and phrases. A pulsar train can last anywhere from a few hundred milliseconds to a minute or more.

Notice in (b) that the *duty ratio* or *d:s ratio* varies while *p* remains constant. In effect, one can simultaneously manipulate both fundamental frequency (the rate of pulsar emission) and what we could call a *formant frequency* (corresponding to the duty cycle), each according to separate envelopes. Lowering the fundamental means increasing *s,* and raising the fundamental means decreasing *s.*

So far, the structure that we have described is similar to a standard impulse generator. Pulsar synthesis generalizes this configuration in several ways. First, it allows the pulsaret *w* to be any waveform.

Let us assume that *w* is a single cycle of a sine wave. From a signal processing point of view, this can be seen as a sine wave that has been limited in time by a rectangular function *v,* which we call the *pulsaret envelope.* An important generalization is that *v* can also be any shape. The envelope *v* has a strong effect on the spectrum of the pulsar train, as the spectrum is the convolution of *v* and *w.*

Keeping *p* and *w* constant and varying *d* on a continuous basis creates the effect of a resonant filter swept across a tone. There is, of course, no filter in this circuit. Rather, the frequency corresponding to the duty cycle *d* appears in the spectrum as a formant peak. By sweeping the frequency of this peak over time, we obtain the sonic equivalent of a time-varying bandpass filter applied to a basic impulse train.


fundFreq min 0 max 50

formFreq1 min 50 max 190

formFreq2 min 0 max 2000

formFreq3 min 210 max 5450

pulseMask min 0 max 1

pan1 min -1 max 1

pan2 min -1 max 1

pan3 min -1 max 1

pulsaret min -1 max 1

amp1 min -1 max 1

amp2 min -1 max 1

amp3 min -1 max 1

pulseEnv min -1 max 1

Pulsar Generator 1.0.01b-debug

Volume 1 ☐ ON

Volume 2 ☐ ON

Volume 3 ☒ ON

☐ Burst Ratio: on: 3 off: 2

Train Duration: 63.1

☒ Loop ☐ Record ☐ MUTE

You have used 454870 PulsarCredits.

Fade Time: 2

FX : RingMod

No File. RingMod is off.

Figure 2. Screen image of the PulsarGeneratot program by Alberto de Campo and Curtis Roads (2001).

Figure 2 shows the control panel of PulsarGenerator. Most parameters are controlled by time-varying envelopes lasting the stipulated length of a pulsar train. The envelopes can be freely copied, pasted, multiplied and scaled and can be imported from any sample file.

<Give demonstration of PulsarGenerator>

Microfiltration and its musical applications: the SweepingQGranulator program

Operations on the micro time scale can be extended beyond synthesis and into the realm of digital signal processing, or digital audio effects. One of the most interesting paths of exploration has been in the domain of *microfiltration.* The basic idea behind microfilteration is to apply filters on the micro time scale. Before my experiments, this had never been tried before.

Microfiltration can be realized in a number of ways. My first experiments date back to 1997. In these experiments, my software would first granulate a sound sample and then apply a bandpass filter on every grain. The filter applied to a given grain had unique characteristics, due to stochastic controls. Thus it was possible to hear the effect of potentially hundreds of different filters per second–a scintillating sound.

The most interesting experiments involved constant *Q* filters. The *Q* of a bandpass filter can be defined as the ratio of the center frequency to the spread of its -3 dB point (cutoff point) bandwidth. Notice that when the center frequency is constant, adjusting the *Q* is the same as adjusting the bandwidth. A constant *Q* filter, on the other hand, adjusts the bandwidth according to the center frequency, keeping the ratio the same. For example, suppose that we set the *Q* to be a constant value of 2. When the center frequency is 250 Hz, the bandwidth is 125 Hz. When the center frequency is 2500 Hz, the bandwidth is 1250 Hz. Constant *Q* filters have the advantage that they sculpt the same musical interval regardless of their center frequency. When the filter *Q* and grain density are both high, the granular stream takes on a liquid quality.

My software SweepingQGranulator lets one control these parameters in real time as the sound is heard, and can also sweep the filter center frequency on every grain. This per-grain filter sweep adds more dynamic timbre variation.

<Give demonstration of SweepingQGranulator>

Pulsar synthesis and microfiltration work well in combination. For example, I have produced the sound material for several compositions, including *Granules* (1999), *Tenth Vortex* (2000) and *Eleventh Vortex* (2001), by combining pulsar generation with microfiltration.

Matching pursuit decomposition

One of the principle barriers to past methods of time-frequency (TF) analysis has been the attainment of high resolution in both time and frequency. In recent years, techniques that exploit so-called *atomic representations* have emerged as useful alternatives. Atomic representations model sound as an agglomeration of TF *atoms*. This paper presents one such method, called *matching pursuit* (MP) analysis (Mallat and Zhang 1993; Mallat 1998), which offers new and interesting ways to view and interact with audio information.

The MP technique assumes that any sound can be decomposed into a combination of atoms predefined in a dictionary of acoustic particles. MP analysis searches to find a correlation between a given region of TF energy and an atom in the predefined dictionary.

Our many experiments, carried out over a period of several years, have shown that this technique leads in promising new directions for audio signal processing. These directions derive from the robust, high resolution atomic representation produced by MP. By robust, we mean that the analysis results in a list of sound atoms. One can, in general, modify or delete a given atom without introducing unwanted audio artifacts. By high resolution, we mean that the TF energy can be resolved to an arbitrary degree of precision. The main constraint is computation time, which is a serious issue with MP analysis.

MP decomposition does not proceed by analyzing successive windows or time frames as to their energy in specific frequency bins. Instead, it operates on the entire signal in the time domain. MP decomposition localizes the energy to a unique atom that can be manipulated independently, making it very useful the development of new digital audio effects.

The visualizations produced by MP decomposition can reveal TF structures that are blurred in the traditional sonogram, particularly transients, finely spaced frequencies, and the detailed structure of complex noises. Novel transformations based on the atomic representation generated by MP analysis extend the artistic toolbox of composers and sound designers.

Let us look at the representation of a speech waveform. In the top, we see the time-domain image; in the middle is a traditional sonogram representation produced by short-time Fourier analysis, and in the bottom display we see the same signal represented by matching pursuit decomposition as an *atomic sonogram.*

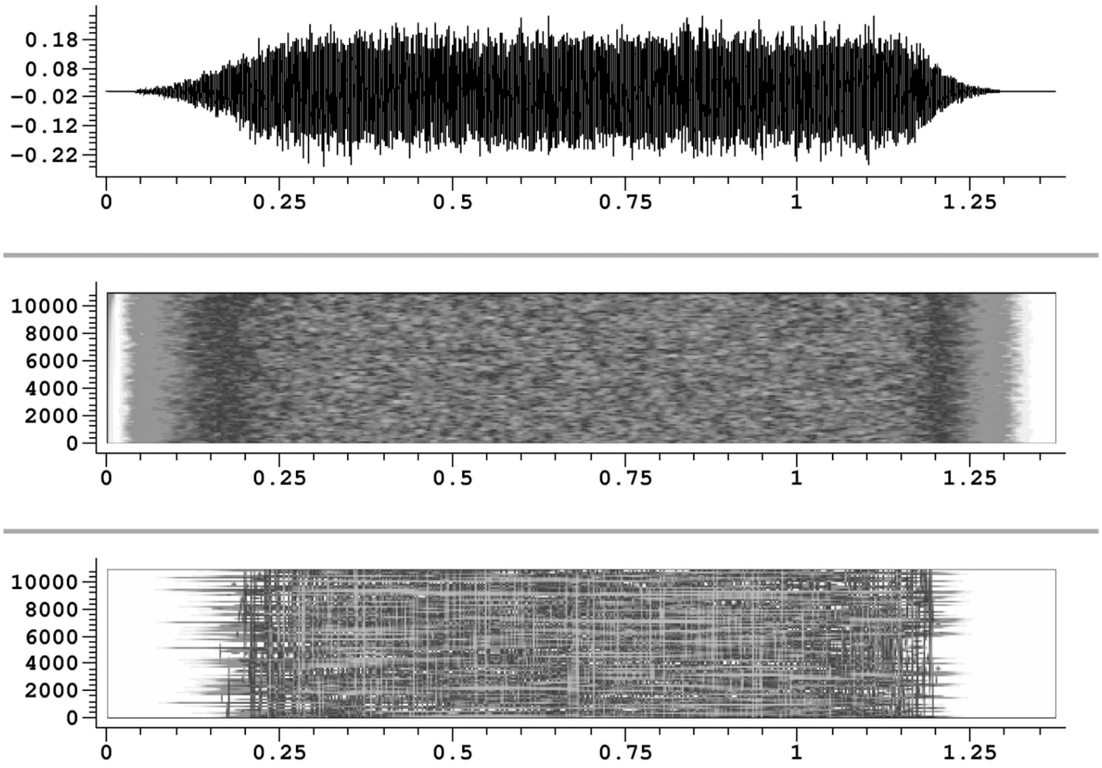


Figure 3. Noise. Top: Time domain waveform. Middle: Short-time Fourier Transform. Bottom: Matching pursuit decomposition.

In Figures 3 and 4, notice the blurriness and noise in the Fourier sonogram (middle), as compared to the precise localization of energy in the atomic sonogram (bottom).

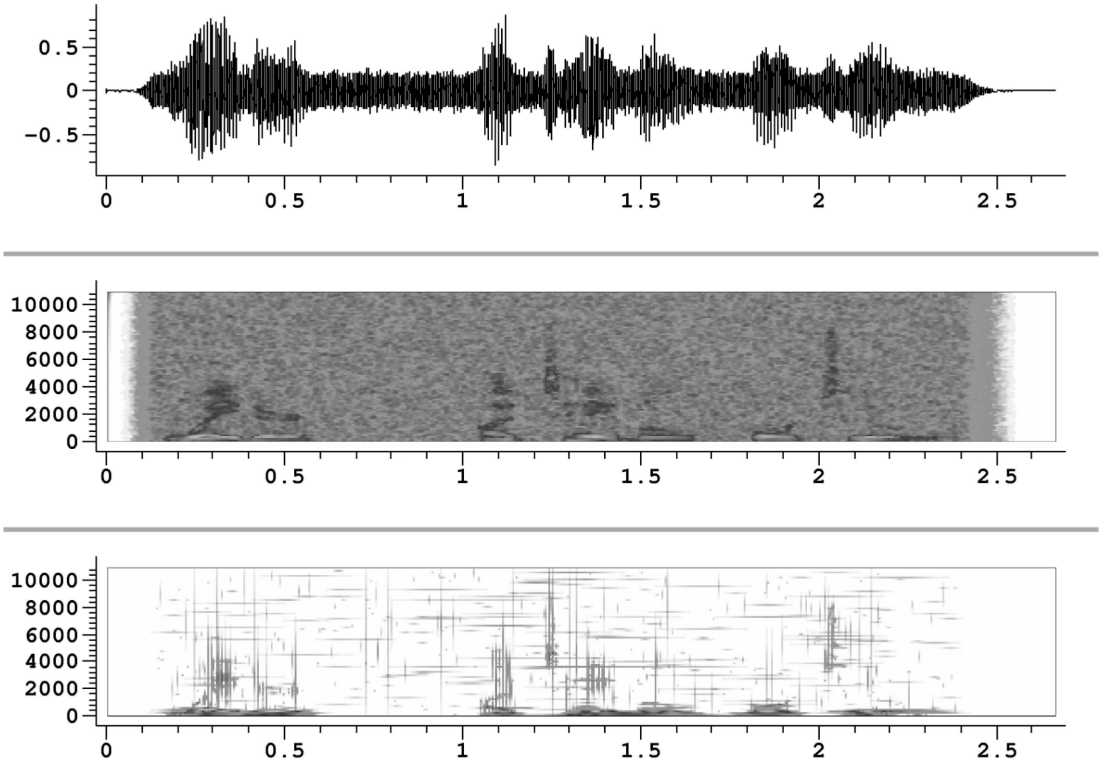


Figure 4. "Waiter, a table for two" + noise. Top: Time domain waveform. Middle: Short-time Fourier Transform. Bottom: Matching pursuit decomposition.

The core of the MP technique is time-domain comparison process based on a brute-force search of a predefined dictionary of wavelet particles. For this reason, it is computationally expensive, with typical compute ratios on the order to 200 to 1 on a single processor. However the algorithm is linearly scaleable so I am confident that it can be highly optimized by an implementation on a computing cluster.

My interest in MP decomposition is based on these qualities:

arbitrarily high resolution - dependent only on dictionary size
robust analysis data - each particle can usually be manipulated more-or-less independently of the other particles

Contrast this with the fragile interdependent data produced by the traditional *short-time Fourier transform* (STFT), for example, where the energy at any instant depends on summing all of the frequency bins, and the borders of each frame must align in phase with preceding and successive frames, so that any modification within a frame is very likely to result in a discontinuity.

By the way, once the analysis is complete, the MP analysis data is highly compressed, on the order of a typical MP3 file, and can be easily resynthesized in real time.

Cavitation processes

Sound particles dissolve the rigid bricks of musical composition–the notes and their intervals–into more fluid and supple materials. We can now shape sonic matter in terms of its particle density and opacity. Particle density has become a prime compositional parameter. Physics defines density as the ratio of mass to volume. In music this translates to the ratio of sound to silence. Through manipulations of density, processes such as coalescence (cloud formation), and evaporation (cloud disintegration) can occur in sonic form.

Opacity correlates to density. If the density of microsonic events is sufficient, the temporal dimension appears to cohere, and one perceives a continuous texture on the sound object level. Thus by controlling the density and size of sound particles we have a handle on the quality of sonic opacity. Coalescence takes place when particle density increases to the point that tone continuity takes hold. An opaque sound tends to block out other sounds that cross into its time-frequency zone.

The aesthetic frontier of granular analysis and synthesis is in cavitation processes. Many physical acousticians consider the phenomenon of sound to be the product of fluid mechanics, as the laws of wave propagation are essentially the same in gasses and liquids. Consider the phenomenon of bubbles in a liquid medium. As we all know, bubbles are sound emitters. Each bubble is a resonator.

We can use cavitation processes as filters to sculpt the internal morphology of a sound on a micro time scale. We can also use cavitation as a control over the birth and death of sound objects, making it possible for sounds to *coalesce* into being, or *disintegrate* into nonbeing. Thus control over the virtual condensation and evaporation of sounds in a further realization of Edgard Varèse's vision

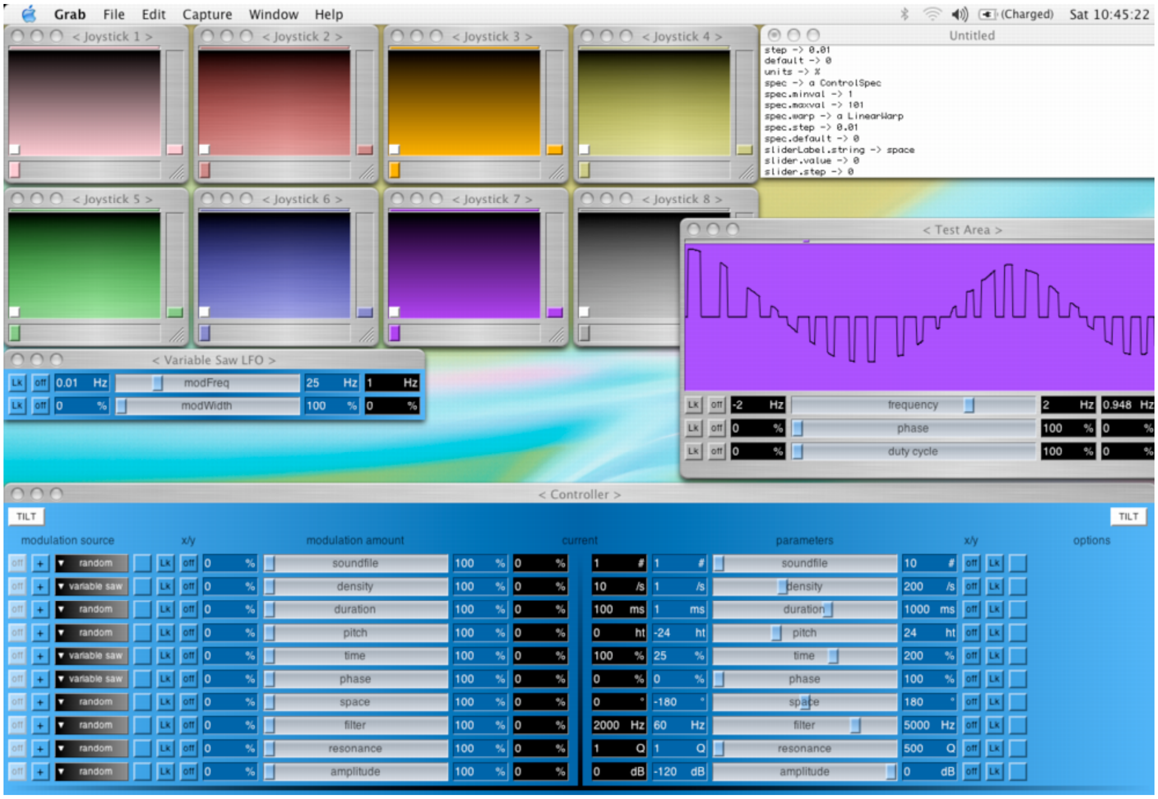
<Play disintegration examples vinoDis3 and 4>

EmissionControl

EmissionControl (Figure 5) is a new program designed by David Thall and myself for generalized particle synthesis and granulation. It features a sophisticated control scheme that is managed by a novel user interface. The user interface helps users to navigate the multidimensional parameter space inherent in granular transformations. The combination of efficient processing with scalable and customized controllers results in a system designed exclusively for advanced granular synthesis and processing. This system has been proven in both studio and live onstage settings, and could also be used in applications such as sonification of scientific data or as a synchronized accompaniment to visual generation.

Granular sampling techniques let a composer stream or scatter acoustic particles in multiple dimensions, either in real-time or by script. Using various granular synthesis and processing models, a set of parametric control data is generated and mapped to an underlying grain-scheduling algorithm. In this context, a sound grain can be considered a finite time segment of an arbitrary waveform modulated (shaped) by an amplitude envelope.

Historically, granular techniques have been used for synthesis and in commercial pitch-time-formant changing effects. Our efforts in the EmissionControl project have focused on widening the range of possible sound enhancements and transformations. The ability to work with sound at the level of individual sound grains opens up the possibility to deconstruct and reassemble sounds anew, into innovative shapes and textures.



```
step -> 0.01
default -> 0
units -> %
spec -> a ControlSpec
spec.minval -> 1
spec.maxval -> 101
spec.warp -> a LinearWarp
spec.step -> 0.01
spec.default -> 0
sliderLabel.string -> space
slider.value -> 0
slider.step -> 0
```

modulation source	x/y	modulation amount			current	parameters		
random	0 %	soundfile	100 %	0 %	1 #	1 #	soundfile	10 #
variable saw	0 %	density	100 %	0 %	10 /s	1 /s	density	200 /s
random	0 %	duration	100 %	0 %	100 ms	1 ms	duration	1000 ms
random	0 %	pitch	100 %	0 %	0 ht	-24 ht	pitch	24 ht
variable saw	0 %	time	100 %	0 %	100 %	25 %	time	200 %
variable saw	0 %	phase	100 %	0 %	0 %	0 %	phase	100 %
random	0 %	space	100 %	0 %	0 °	-180 °	space	180 °
random	0 %	filter	100 %	0 %	2000 Hz	60 Hz	filter	5000 Hz
random	0 %	resonance	100 %	0 %	1 Q	1 Q	resonance	500 Q
random	0 %	amplitude	100 %	0 %	0 dB	-120 dB	amplitude	0 dB

Figure 5. Screen shot of EmissionControl (2005).

<Demonstrate EmissionControl and play demonstration of EmissionControl sounds>

Granular visualizations

Another thrust of research at our laboratories falls under the rubric of visualization and sonification; this is the Ynez research project. Ynez stands at the nexus of a number of overlapping research problems in the relations between sound, image, and music notation:

- Interactive composing environments based on drawing and manipulation of images of waveforms, envelopes, and sonographic spectra
- Scientific visualizations of sonic data based on sound analysis
- Scientific sonification of image (or other data) data to sound
- Artistic visualizations of music, either abstract or representational music animations
- Study scores and for electronic music, comprising still images that intermingle sonographic, iconic, and symbolic representations

Technological advances have accelerated efforts to visualize and sonify. Software translations between sound and image that were once the province of laboratory specialists are now accessible to anyone with a computer. As a result, activity in all these intertwined areas is rapidly evolving.

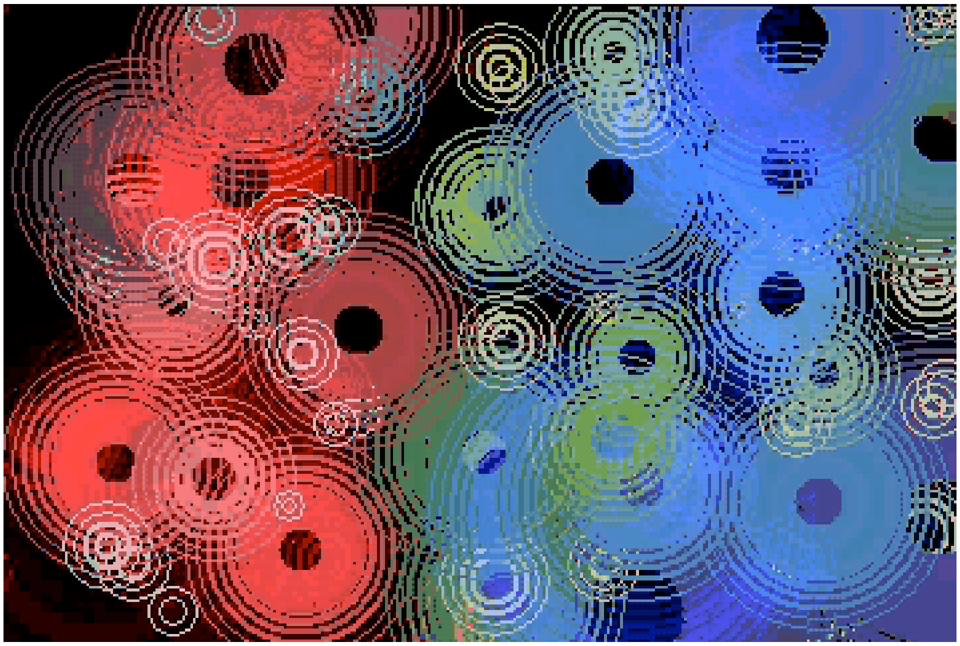


Figure 6. *Sonal atoms* by Curtis Roads, scientific visualization by Woon Seung Yeo.

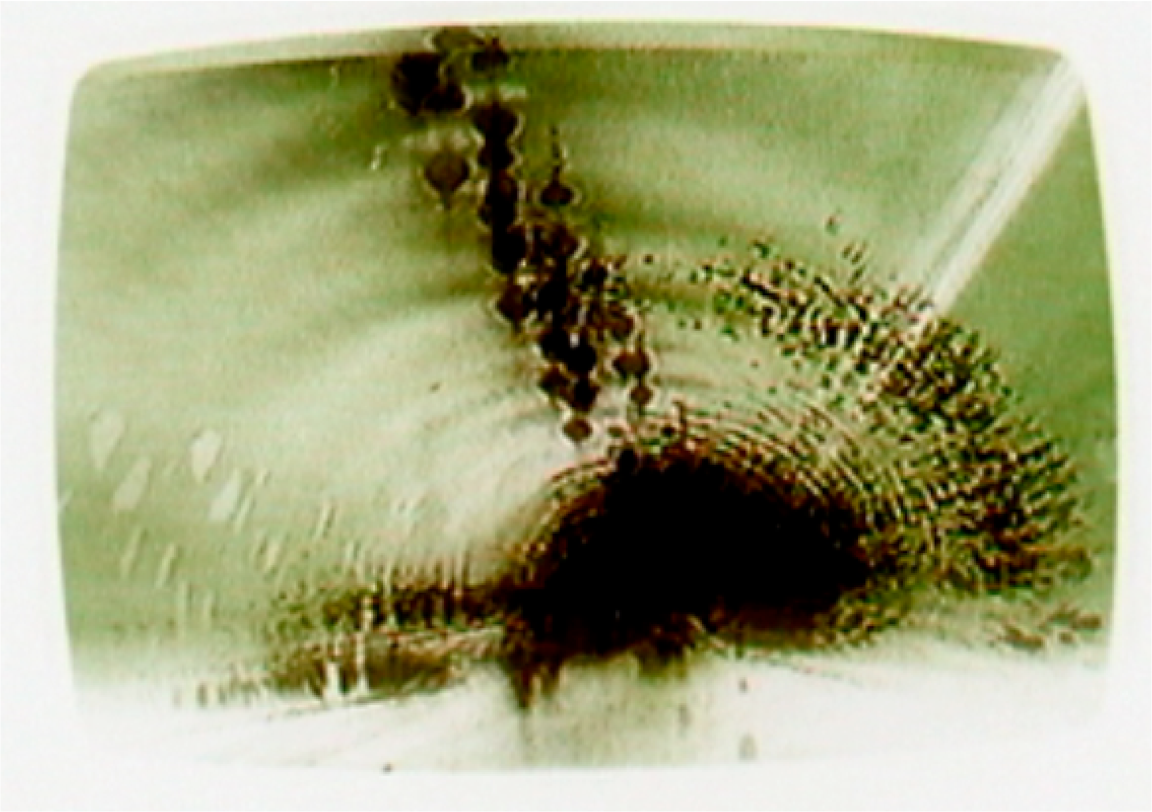


Figure 7. *Sonal atoms* by Curtis Roads, artistic visualization by Brian O'Reilly.

Figures 6 and 7 show two visualizations of the same piece, one scientific (a precise mapping of time-frequency energy into a two-dimensional image) and one artistic (video feedback edited to synchronize to the music).

Conclusion

> *If there is one dimension of the music totality, one component which originally led composers to the electronic medium, it was and is the temporal domain. ... Those who originally turned to electronic tape were obviously attracted to the element of control. After all, the tape was not a source of sound. Tape is for storage. You can, however, control time as a measurable distance of tape. Here we are talking about rhythm in every sense of the word. Not only durational rhythm, but also the time rate of changes of register, of timbre, of volume, and of those many musical dimensions that were unforeseen until we tried to find out how we heard and how we could structure the temporal.* – Babbitt (1988)

One goal of technology is increased precision and ever-finer control. The trend toward precision is also reflected in our sound tools, which have passed from wave-oriented to particle-oriented operation, even as the sample grid has shrunk to just over 5 μsec (at a 192 kHz sampling rate).

By now, hundreds of compositions employ granular synthesis and other particle techniques. The raw material of this music consists of grains and globules scattered in sonically transparent, diaphanous, or opaque textures. These textures can be molded, stretched, and combined with others into supple morphologies. By means of these materials, musical development can take place not only in the large intervals of pitch and rhythm, but also in microvariations of amplitude, duration, timbre, density, and spatial position.

Perhaps more important than the particles themselves are the sonic brushes we use to scatter them on the canvas of time and frequency. These brushes are computer programs. By connecting these programs with interactive real-time controllers, we have built particle synthesis instruments for virtuoso performance, not only for onstage use, but also for use in the studio.

In the domain of computer graphics, three-dimensional animation programs incorporate sophisticated algorithms for scattering particles. These emulate physical models of flow, perturbation, and collisions among particles. In the domain of sound, we can also apply physical models to regulate the flow of particles. But we should not be limited to emulations of reality. Indeed, the computer's artistic power derives from its ability to model fantasies as well as reality.

Creating sonic fantasies began with recording, letting us "photograph" real sounds and store their images on tape. The techniques of montage–cutting, splicing, and mixing–were essentially manipulations of time structure. Today, discs and semiconductors have largely superseded tape. Mixing and editing software has replaced the splicing block. A fundamental capability of this software is zooming in and out on multiple time scales. These tools let us work at the limits of auditory phenomena, from microsurgery on individual sample points, to the global rearrangement of sound masses.

Acknowledgments

I would like to thank my partners in research: Alberto de Campo, Garry Kling, David Thall, Brian O'Reilly, and Bob Sturm.

References

Gabor, D. 1946. "Theory of communication." *Journal of the Institute of Electrical Engineers* Part III, 93: 429-457.

Kling, G., and C. Roads. 2004. “Audio analysis, visualization, and transformation with the matching pursuit algorithm.” In G. Evangelista, ed. *Proceedings of DAFX 2004.* Naples: Federico Due University of Naples.

Kling, G., B. Sturm, and C. Roads. 2005. “Sound visualization and transformation by means of matching pursuit decomposition.” Submitted to the *Journal of the Audio Engineering Society.*

S. Mallat, *A Wavelet Tour of Signal Processing*. San Diego: Academic Press, 1998.

S. Mallat and Z. Zhang. “Matching pursuit with time-frequency dictionaries.” *IEEE Trans. Signal Proc.*, 41(12):3397--3414, 1993.

Roads, C. 2001. “Sound composition with pulsars.” *Journal of the Audio Engineering Society* 49(3): 134-147.

Roads, C. 2002. *Microsound.* Cambridge, Massachusetts: The MIT Press.

Roads, C., and J. Kuchera-Morin. 2000. “The Ynez notation project.” Unpublished white paper.

Xenakis, I. 1960. "Elements of stochastic music." *Gravensaner Blätter* 18: 84-105.