

Transmission Two: The Great Excursion (TT:TGE)—The Aesthetic, Art and Science of a Composition for Radio

Larry Austin,
Charles Boone and
Xavier Serra

TT:TGE AS HYBRID ART— A LISTENER'S PERSPECTIVE

Charles Boone

HÖRSPIEL

Sonic composition conceived to be heard through loudspeakers is one of the unique and original art forms of the second half of the twentieth century. Concert halls were assumed at the beginning to be the proper venues for presenting this kind of work, but it was not long before composers began searching for other options. Morton Subotnick, for instance, was among the first to try subverting such traditional presentation by composing an electronic work (*Silver Apples of the Moon* [1967]) that would be available only on a long-playing phonograph recording [1]. This piece was intended to be played on a normal high-fidelity system in listeners' homes and in no other way.

In the past decade another idea for bringing original sonic art directly into the home—again, work intended for only this means of dissemination—has begun to flourish. This is *Hörspiel*, a sonic art form that has developed into something truly indigenous to radio. *Hörspiel*, or 'listen play', shortened from the term *Neues Hörspiel*, was coined by German radio producer Klaus Schoening, who, since the late 1960s, has created radio plays by combining sounds and words in nontraditional collages. Schoening has since broadened the concept, now calling it *ars acustica*. It is, as he says, "for ears and imagination" [2].

As the name suggests, there has been a particular flurry of such activity in Germany, but there are now composers and other artists involved with sound and theater who are working on similar projects all over the world.

TT:TGE

In terms of its forces and duration, *Transmission Two: The Great Excursion (TT:TGE)* (1988–1990) may well be the most ambitious *Hörspiel* written so far. It is somewhat of a hybrid case, however, in that it must have a concert performance, with all the preparation that implies, in order for the radio transmission to be possible. This necessity for a concert performance of something really intended for radio presents a certain ambiguity: listeners in the hall might well wonder if all the sounds they are hearing—voices, instru-

ments, computer sounds, pre-recorded texts—are being heard in the same way by radio listeners, and wonder especially whether the sound mixes and balances are the same. One can imagine audience members bringing along portable radios, like spectators at a ball game, just to find out what the outside world is, in fact, hearing. A comparison of a recording of the live performance in Denton, Texas, with a recording taken from its radio broadcast showed considerable differences, particularly in the sound levels of the texts. In the radio mix these sound levels were significantly higher than in the concert hall mix.

DIALOGUE AS PORTRAITURE

The question of balance among the various materials, then, is extremely important, and the levels of perceptibility seem to have been a key concern for Austin in assembling the layers of sound. Of course, the recorded dialogues were central to this consideration since they are the basic building blocks for the entire piece. Because of the nature of the dialogues, *TT:TGE* has a fundamental relationship with documentary reportage, even though the piece is far removed from any traditional work in that genre.

In the final piece, the words operate on two levels: the original, purely informational pre-recorded conversations (what I call 'texts') exactly as they were spoken; and the fragments of these same texts that Austin processed at Stanford University's Center for Computer Research in Music and Acoustics (CCRMA). Obviously, these processed fragments (Austin's 'sonic furies') were pushed well beyond comprehensibility into the realm of pure sound. Perhaps not so obviously, the unaltered texts often function similarly

ABSTRACT

The three authors each explore an aspect of Larry Austin's 1990 extended composition for radio broadcast. Composer-writer Charles Boone examines the work for its musical and dramatic interest and for its significance as a hybrid of *Hörspiel*, with elements of theater, portraiture and sound collage. Its scope, means, intent, realization and compositional approach are elaborated by Austin, the composer. Throughout his section, he intersperses aphorisms drawn from his conversations with the seven protagonists of his 'sound movie'. Austin's collaborator, computer-music scientist Xavier Serra, explicates the technical/theoretical genesis of *TT:TGE*'s computer-generated speech sounds.

Larry Austin (composer, writer, teacher), 2109 Woodbrook, Denton, TX 76205, U.S.A.

Charles Boone (composer, writer), 2325 Bear Gulch Road, Woodside, CA 94062, U.S.A.

Xavier Serra (acoustician, musician), Center for Computer Research in Music and Acoustics, Dept. of Music, Stanford University, Stanford, CA 94305, U.S.A.

Received 12 September 1990.

in that they are not always above the threshold of comprehension; sometimes they are crisp and clear (particularly at beginnings of episodes), and at other times they blend into the background while other musical events take precedence. In traditional documentary such license with a text would be unthinkable.

Why, then, did Austin go to the trouble of recording all these conversations and then use them in ways that often preclude the possibility of their being fully heard and comprehended? When we understand that his intention in *TT:TGE* goes well beyond the simple transmission of information, we can begin to see this work not so much as a documentary, didactic effort but rather as a series of composer portraits on a more abstract level—a musical object based on the words, ideas and personas of a particular group of composers.

THE EPISODES

The interaction of the unaltered recorded conversations and computer-processed sonic furies with the voices of the chorus and the improvised instrumental music of the score is a major element in these portraits. A good example of this amalgamation of diverse components is heard in the Robert Ashley Episode, “The ‘I actually wanted to buy’ Blues”. The movement begins with keyboard and percussion music. The keyboardist improvises in a slow-blues style reminiscent of music in Ashley’s video opera, *Perfect Lives* [3]. (The harmonic structure has been laid out for each part, and audiotapes of slow-blues-style music that model the intended effect have been provided to each performer.) Likewise, the percussionist improvises a jazz accompaniment using wire brushes on suspended cymbals. These musicians are soon joined by the celletto (electric cello) player, who plays fully notated, blues-inspired licks, and by chorus members, joining in one by one, singing and speaking. By this time, recorded dialogue between Ashley and Austin is also playing and continues for the full duration of the movement. The texture builds and becomes more complex, leading to a high point three quarters of the way through the piece, where all chorus members are singing in unison. The movement ends with a whispered section of eight independent chorus parts and continued instrumental improvisations. The choral materials complement the computer-music sounds with which they are layered in this closing section.

As this brief description shows, the movement has a clearly developed form. The combination of musical elements with the prerecorded conversations from which these elements were derived is carefully conceived and worked out to heighten the musical message as well as to make the ‘picture’ of Ashley somehow a true one. The mood and spirit of this movement come directly from Ashley’s own work; there is a very relaxed quality, studied in its casualness, that anyone familiar with his music will recognize immediately. The general effect of the whole episode is almost like sitting in a crowded bar with a blues piano playing in the background, snippets of many conversations drifting in from other booths (sometimes heard and sometimes only perceived), occasional laughter, and a gentle barrage of group noise and other ambient sounds. This is a clever musical portrait that transcends the spoken text that was the starting point for its composition.

All of the episodes are similarly executed, using all of the elements—and especially the choral and instrumental elements—to reflect some of the distinctive characteristics of each of Austin’s subjects. The use of textural material in *TT:TGE* is found in a number of Austin’s earlier works as well. His process, here and elsewhere, is generally to provide materials (on tape, for players, etc.) that he then literally and/or figuratively ‘switches on’ and lets run, allowing them to interact and connect, however serendipitously. A comparison of the first two episodes—the John Chowning and the Jerry Hunt—clarifies this and sheds light on the organization of the other sections as well.

The John Chowning Episode is in three parts (A B A’), which one can easily see by examining the chorus part separately. The first and last of these parts (A and A’) are essentially vertical and harmonic in style, while the long middle part (B) is linear and contrapuntal. The B section is canonic in form, with the imitative voices slowly entering in pairs (sopranos, followed by altos, then tenors, then basses), thus thickening the contrapuntal texture and building over a period of about 5 min to the piece’s first major high point. With these voices are layered the spoken text and the sonic furies, plus the instruments, which provide a long, sustained droning sound and rich coloration. One is not so much aware of the individual voices—particularly of

the chorus’s contrapuntal lines—as of the growth of the texture and the progression from particularity to generality. Perhaps, the traditional A B A’ form and the use of strict contrapuntal procedures somehow reflect the academic setting of CCRMA, where Chowning does his work. This, however, is only speculation.

The Jerry Hunt Episode, which follows on the heels of the Chowning, provides an abrupt and lively contrast. It is a leap from the structure of academia to the freedom and chaos of the world of rap music. Here, there is the driving beat of a drum machine from the first measure to the last, in contrast to the absence of pulse that characterizes the previous episode. The choral singing of the Chowning Episode is replaced with choral speaking. The singers are assigned rhythmic patterns that match fragments of Hunt’s spoken text, which are repeated over and over for long stretches of time. The clarity of the John Chowning Episode gives way here to a wild free-for-all that mounts to a blast of noise. This steady-state sonic mass does include some variety as the movement goes on, but essentially, it can be heard as a large wave of sound that rolls along until it suddenly diminishes at the end, like a phonograph that is switched off while the needle is still in the groove. The text in this movement is almost never heard well enough to be understood. Nonetheless, it is Hunt’s amusing and lively monologue (it can be read in the score) that provides the overall character of the music, and the quality of his rapid banter clearly influenced the way in which the choral and instrumental parts were conceived.

The remaining four episodes also have their own special formal characteristics, use of materials and modes of portraiture. The David Tudor Episode features a tango, in its instrumental and vocal parts, that is heard throughout, as well as rain sounds in the electronic part that relate to Tudor’s own composition, *Rainforest I–IV* (1968–1973). For the most part, the chorus sings, but there are whispered sections as well. The Max Mathews Episode is the gentlest of the seven, perhaps reflecting that aspect of the subject’s character. For the larger part of the movement the men of the chorus whisper, while the women sing. This all begins in fragmentary fashion, but as the movement progresses, the materials become more coherent and coalesce into the late nineteenth-century song “Daisy Bell”,

ending on a full-blown F-major chord. In the Morton Subotnick Episode, the chorus begins whispering in unison, which evolves into spoken polyphony near the end. The Pauline Oliveros Episode is marked by chordal choral writing throughout, with the chorus singing and sometimes humming as well. Near its end the audience is invited to join in by humming—a kind of meditational participation that comes directly from Oliveros's practice of involving listeners in her own performances.

CONCLUSION

Austin's *TT:TGE* is an assemblage of a very large number of materials, ranging from traditional instrumental and vocal practices and straight-spoken texts to innovative live-performance practices and computer-generated sounds. By using recorded dialogues (more normally encountered in a documentary context) to generate abstract and purely musical materials, Austin creates a type of historic capsule of some of the composers and music of our time, while simultaneously interpreting and commenting on the materials therein. Perhaps this fresh approach to portraiture and documentation is more informative (as well as being more entertaining) in the long run than if Austin had simply laid out the facts in a more straightforward but less original manner.

MODELING *TT:TGE*—THE COMPOSER'S PERSPECTIVE

Larry Austin

A RADIO PIECE

I compose computer music. . . . It was absolutely an ear discovery. . . . Music is as abstract as any mode of art. . . . Push the technology, push the technology!

—John Chowning

The premiere of my composition for radio, *Transmission Two: The Great Excursion (TT:TGE)*, was broadcast live in a public performance from the Berkeley campus of the University of California [4]. *TT:TGE* is scored for a chorus of 40 voices (soprano 1–2, alto 1–2, tenor 1–2, bass 1–2); a computer-music ensemble, including keyboardist, cellist and percussionist, with technical assis-

tants; and prerecorded dialogues. The dialogues heard throughout the work present a chronicle of my conversations, recorded between November 1988 and July 1989, with fellow composers Robert Ashley, John Chowning, Jerry Hunt, Pauline Oliveros, Morton Subotnik and David Tudor and computer-music pioneer Max Mathews. With me, they reflect on their work as composers and music experimenters today and through the last 30 years of dramatic technological developments in the way music is created and presented.

FORM AND CONTINUITY

I manipulate convention . . . the expectations of performance and the memory and the comparisons and ideas of performance and the relationship to, uh, to the gestures that are associated with performance. . . . I decided to compose with that . . . by getting rid of the idea of music itself and just going directly to just memory and retrieval and history of everything you've done. . . . Then, you can carry, kind of in a way . . . it's a little bit like exchanging cassettes in a social . . . in a socially charged atmosphere. . . . I have no feelings about taste or anything. . . . He's like a Mozart . . . all the conventions, he . . . accepts them all. I mean they're all there: the clichés about sex and power and black and white and money and everything and the chords and the rhythm machines. . . . I've always cheated. . . . What do you want from me now? It's kinda' like God, you know, is a puppeteer up there, you know, and we're all playing down below, and this machinery is just grinding . . . but I understand it better, as I'm getting older myself because, you can hear that coffin snappin', you know! And you're thinkin', "Snap, snap, snap, snap . . . it's snappin' for me." And when you hear that snap, you think, "Do . . . ha . . . how . . . you know, I want a good snapshot, I want a good snapshot left."

—Jerry Hunt

The first live broadcasts of *TT:TGE* culminated 2 years of planning, recording, modeling, computing, experimenting with and composing the materials for what I have come to call my 'sound movie'.

Central to the form and narrative

continuity of *TT:TGE* are seven episodes, each devoted to a dialogue between me and one of my longtime compatriots—adventurers with distinctively etched musical and technological profiles. Key segments from more than 11 hours of these recorded conversations were distilled to 107 min: the Prelude and Episodes 1–4 (Chowning, Hunt, Ashley, Oliveros) form Part I; Episodes 5–7 (Tudor, Mathews and Subotnik) and the Postlude form Part II. The episodes range in duration from 12 to 16 min each.

The musical continuity of *TT:TGE* unfolds through the episodes with choral settings and computer-music transformations of select aphorisms uttered spontaneously during the conversations I recorded—essences, I feel, of the protagonists' aesthetic and compositional approaches. The role of the chorus in my sound movie is as choric commentator, not unlike its role in classical Greek theater: musing poetically, reflecting philosophically, interjecting assertively, commenting amusedly or mocking 'rap'-ingly.

APHORISM AS SONIC MATERIAL: TECHNICAL REALIZATION

We studied all those forms in school. . . . They were the models for us, and we had no contact with them at all. . . . There's no way that I could be comfortable writing a sonata. . . . I mean that's extremely sentimental . . . because the language doesn't contain any words . . . mechanical marvels, magic . . . mystery machines. . . . I felt this growing excitement in me . . . because I hadn't bought anything . . . almost like being in love. It was almost like sex. I felt the great desire to own a new piece of equipment. I hadn't owned anything. I became like the most avid consumer. I actually wanted to buy.

—Robert Ashley

The tens of aphorisms selected from the seven conversations created rich material for a wide range—literally hundreds—of sonic transformations. Each aphoristic utterance was first analyzed, then transformed and resynthesized with a unique computer-music software system called Sound Analysis System (SANSY), recently developed for sound analysis/transformation/resynthesis. Named Spectral Modeling Syn-

thesis (SMS), it was developed by computer-music scientist Xavier Serra at Stanford University's Center for Computer Research in Music and Acoustics (CCRMA) [5]. While this experimental system was being perfected and tested in 1989, I was invited by CCRMA Director John Chowning and Serra to explore the compositional potential of SANSY in the creation of sonic materials for *TT:TGE*. After an intensive introduction to the complexities of SANSY operation, I undertook a virtual marathon of intensive experiments during my composer residency there from May to July of 1989 and again in December of 1989, when I completed the final sonic transformations for *TT:TGE*. I can state that SANSY is one of the most powerful tools for sonic transformation I have used in my work as a computer-music practitioner over the last 22 years.

These transformed sonic apherisms—what I call my 'sonic furies'—are heard as musical events in diverse timbral/textural combinations, invoked through the piece by three musicians of a computer-music ensemble with unique computer-music instruments: 'controllers' of multiple Sequential Prophet-3000, 16-bit digital sound 'samplers' (a specialized type of digital recorder) [6]. Through programming the protocol of Musical Interface Digital Interface (MIDI) interaction between the controllers and 40 banks of over 100 'presets' of sonic events stored on a cartridge hard-disk drive for fast access by the multiple samplers, the musicians play and layer my sonic furies through the course of the piece. Typically, the furies are as short as a compressed 1-sec utterance (e.g. "space to sing by itself", uttered by David Tudor) or as long as a stretched 32-sec utterance (e.g. "exactly

from listening, exactly in the moment", quietly spoken by Pauline Oliveros).

PITCH AND DURATIONAL MAPPING

... exactly from listening, exactly in the moment ... to experience the body as a resonator ... the body as a resonator of sound. ... Hear everything that's going on ... be open to the sounds ... deep listening ... becoming a part of what you're listening to. ... That was a nice moment.

—Pauline Oliveros

The rhythmic, melodic, harmonic and textural material for much of the choral and ensemble music was extrapolated from the durational and harmonic partial data yielded from the Fourier

Pauline Oliveros Episode: *Madrigale profondo nella terza prattica*

Fig. 1. Extrapolated pitches/durations from Fourier transform of Pauline Oliveros's uttering 'deep listening'.

Table 1. Vertical pitch/durational combinations extrapolated from frequency of vocal formant.

Syllable:	deep	lis-	t'ning	[sigh...]
Begin-time:	0"	.225"	.35"	.725"-1.175"
Hertz=pitch	195=G3	196=G3	219=A3	[noise]
	481=B4	392=G4	403=G4	[noise]
	585=D5	588=D5	638=D#5	[noise]
	978=B5	782=G5	752=F#5	[noise]
	1175=D6	981=B5	972=B5	[noise]
	1370=F6	1176=D6	1036=C6	[noise]
	1761=A6	1582=G6	1562=G6	[noise]
	1954=B6	1965=B6	1907=A#6	[noise]

transform and the subsequent analyses provided in SANSY's graphic and numerical representations of harmonic and amplitude content of the sampled utterances over time [7]. Figure 1 illustrates this process in the opening of the Pauline Oliveros Episode.

The utterance "deep listening" is stretched 10 times its original length of 1.175 sec to a duration of 11.75 sec. The pitches for the eight-part choral sonorities were then extrapolated from the first eight partials of the deterministic analysis, mapping each frequency to its nearest 12-tone-per-octave equal-tempered pitch, then transporting an octave lower to the choral vocal range, shown in Table 1.

I am particularly attracted to the timbral fusion achieved when the choral sonorities and the matching sonic fury are heard in rhythmic/harmonic unison. Many of the other transformed utterances—especially those where the choral setting called for non-pitched, speech-like execution—were mapped 'by ear'.

COMPOSITIONAL APPROACH

... this dream-vision of an orchestra of loudspeakers ... those are real physical instruments. ... They became my friends. ... It's an act of discovery. ... Release what's there. ... I don't compose in a straight line. I always compose around the periphery. ... I circle it all the time. ... the boundaries, dealing with the edges ... transformations ... space to sing by itself.

—David Tudor

Three things run through my work as a composer and all are present in every detail of *TT:TGE*. First—and most important—I believe that the musics of humankind are concomitant with a plu-

ralist world, one where human rights are freely and creatively experienced and expressed. With that idealized reference, I experience and assimilate all kinds of music. But, in spite of my appetite for tasting the fruits of diverse musical expressions, the influence they have on my own work is elusive and perhaps not one that a composer should so self-consciously contemplate. I believe there is a process of synthesis taking place inside me among these musics I enjoy and assimilate. In *TT:TGE* this became a subtle but integral part of my modeling process, a kind of hybrid synthesis of my assimilation of my subjects' musics, their compositional attitudes and approaches, and my own.

I think the person who has the fun is the performer.

—Max Mathews

Second, I am intrigued by the results of juxtaposing seemingly disparate compositional processes, materials and media, exemplified in much of my work and certainly in *TT:TGE*: the juxtaposition of sonorous choral homophony and strict canonic procedure in the John Chowning Episode; of nearly unperformable rhythmic complexity with a sound-mass thicket of event patternings in the Jerry Hunt Episode; of paradox and languor in the Robert Ashley Episode; of quiet spirituality and lively interaction in the Pauline Oliveros Episode; of the prosaic and enigmatic in the David Tudor Episode; of simple pleasures and inventive genius in the Max Mathews Episode; of medium and message in the Morton Subotnick Episode. For myself as a composer, there is a doubled duality of sentimentality and abstraction contradicted by irony and corporeality. Dualities fascinate me because they are never really understood, just as the polarities in the fortunes of life are never fully reconciled.

Organic growth was the nature of the medium for me. ... We all grew up under the sign of McLuhan ... medium and message. ... It's the gesture. ... The color is the language of the gesture. The gesture is the motive, and the color is the language, just as words—pitches—are the words of language. ... I have to mean something, and if I mean something, it will be colorful. ... There was a hippopotamus, a purple hippopotamus that was the sink in the bathtub, and it occurred to me that the first thing I should do with a big computer is to turn it into a purple hippopotamus. ... I'm the master of the medium. ... Technology has to make the meaningfulness more meaningful. ... If we lived in a world where the butterfly was a metaphor, how would politics be, how would education be? I wanted it to work globally, so that it could work individually for a piece of music.

—Morton Subotnick

Third—and finally—I am happiest in my work when I am exploring fresh concepts and new materials. If my explorations take me to a rich vein of musical material, I experiment intensely with the possibilities. The material instructs me in how to form the piece, or even a series of pieces with kindred properties. I learn from all, assimilating the musical implications. I have also been gratified by the new musical potentialities I have discovered for my pieces to take root in and by the sometimes surprisingly successful (to my ear) musical inventions that grow from my experiments. Always at the base of this approach is the essence of the piece I am imagining. My explorations and experiments with technology and computer-mediated approaches come about for musical reasons, not stemming so much from the newest development in music technology for its own sake as from musical imperatives that—it turns out—often cannot be achieved without advanced technology to use as my *modus operandi*. Thus, I am uncomfortable when I am sometimes introduced as an 'electronic composer'. I am simply a composer of music, not an organism interfaced with a machine. On the other hand, that is not necessarily a flawed concept. I wonder ...

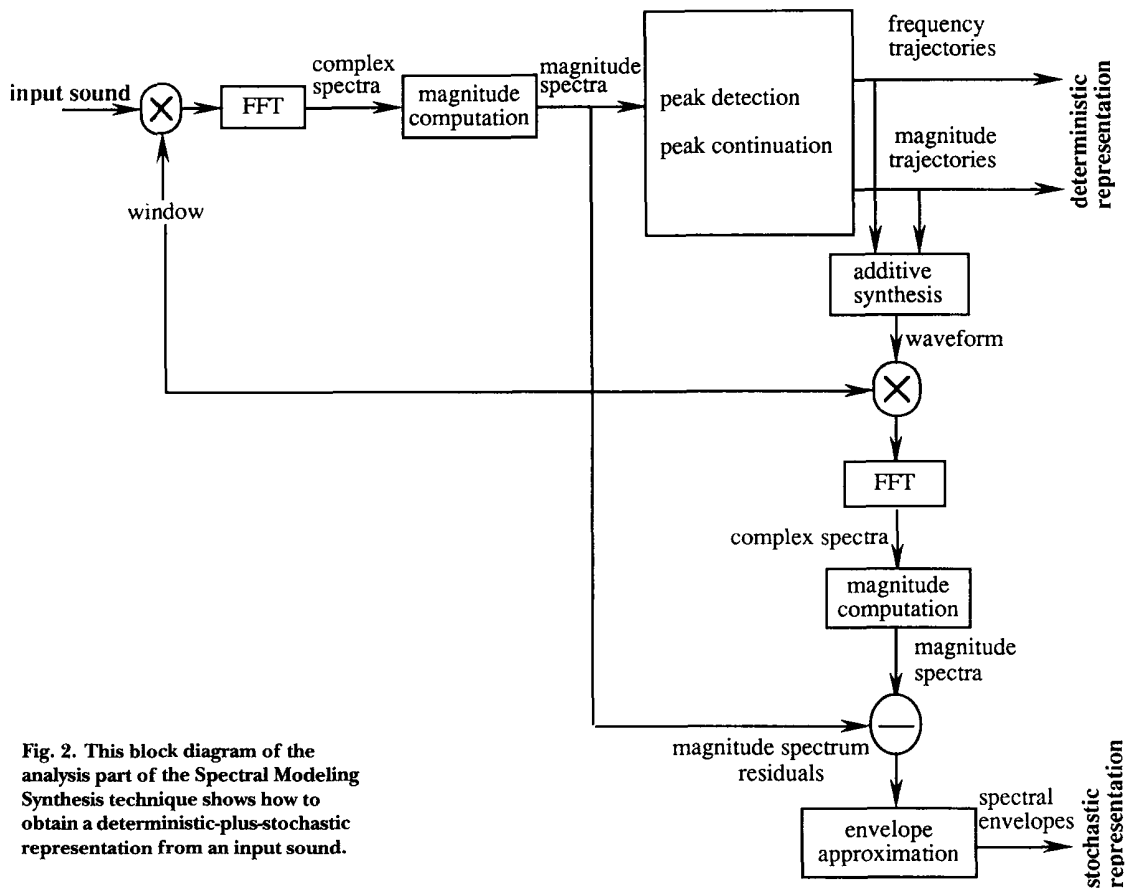


Fig. 2. This block diagram of the analysis part of the Spectral Modeling Synthesis technique shows how to obtain a deterministic-plus-stochastic representation from an input sound.

TT:TGE AND SANSY—THE SCIENTIST’S PERSPECTIVE

Xavier Serra

The piece *TT:TGE* by Larry Austin was the first composition to make use of the Spectral Modeling Synthesis (SMS) technique and of the program SANSY [8]. This computer system, which I developed at CCRMA, allows for the manipulation of prerecorded sounds and was used by Austin to transform a series of recorded voices into his sonic furies.

SANSY is a software environment written solely for the purpose of developing the SMS technique. SANSY is considered more a research tool than a practical program for musical applications. Using such a technically daunting program to create music poses a special challenge to the composer.

SPECTRAL MODELING SYNTHESIS

When generating musical sounds on a digital computer, it is important to have a good model with parameters that provide a rich source of meaningful sound transformation. Good examples of parameters are pitch, amplitude, brightness, and bow pressure. Three basic model types are in prevalent use today: instrument models, spectrum models and abstract models. Instrument models attempt to characterize the parameters of a sound at its source, which might be a violin, clarinet or vocal tract. Spectrum models attempt to describe the parameters of a sound at the basilar membrane of the ear, discarding information from the sound spectrum that the ear seems to discard. Abstract models, such as Frequency Modulation (FM) (the technique used in the Yamaha DX7 synthesizers), attempt to provide musically useful parameters in a mathematical formula.

SMS models sound at the spectrum level. It models time-varying spectra as (1) a collection of sinusoids, controlled through time by piecewise linear amplitude and frequency envelopes (the ‘de-

terministic’ part) and (2) a time-varying filtered noise component (the ‘stochastic’ part). The SMS technique has proved to give general, high-quality transformations for a wide variety of musical signals.

The Deterministic-Plus-Stochastic Model

The SMS technique assumes that the input sound to be transformed has a deterministic component and a stochastic component. A deterministic signal is traditionally defined as anything that is not noise (i.e. a perfectly predictable part, predictable from measurements over any continuous interval). However, in our model the class of deterministic signals considered is restricted to sums of sinusoidal components. Each sinusoid models an actual sinusoidal component of the original sound and is described by an amplitude and a frequency function. A stochastic, or noise, signal is fully described by its amplitude and its general frequency characteristics. When a signal is assumed stochastic, it is not necessary to preserve either the instantaneous phase or the exact frequency information.

Therefore, the input sound $s(t)$ is the sum of a series of sinusoids plus a noise signal,

$$s(t) = \sum_{r=1}^R A_r(t) \cos[\theta_r(t)] + e(t) \quad (1)$$

where $A_r(t)$ and $\theta_r(t)$ are, respectively, the instantaneous amplitude and the phase of each sinusoid and $e(t)$ is the noise component.

The model assumes that the sinusoids are stable partials of the sound and that each one can be characterized by its amplitude and frequency. The instantaneous phase is then taken to be the integral of the instantaneous frequency $\omega_r(t)$, and therefore satisfies

$$\theta_r(t) = \int_0^t \omega_r(\tau) d\tau \quad (2)$$

where $\omega(t)$ is the frequency in radians, and r is the sinusoid number.

By assuming that $e(t)$ is a stochastic signal, it can be described as filtered white noise,

$$e(t) = \int_0^t h(t, \tau) u(\tau) d\tau \quad (3)$$

where $u(t)$ is white noise and $h(t, \tau)$ is the impulse response of a slowly time-varying filter. That is, the residual is modeled by the convolution of white noise with a frequency-shaping filter.

The filtering of a noise signal can be implemented by multiplying the inverse Fourier transform of the filter-frequency response by a random phase term. This method of filtering a noise signal is used to synthesize the stochastic signal.

Description of the SMS Technique

Figures 2 and 3 show the block diagrams for the analysis and synthesis parts of the SMS technique. The first step in the analysis is the derivation of a series of magnitude spectra of the input sound by computing the Fast Fourier Transform (FFT) of every windowed portion of the waveform, i.e. computation of the Short-Time Fourier Transform (STFT). Both the FFT and STFT are mathematical operations that can convert a portion of a sound into frequency spectra. From the series of magnitude spectra the analysis detects prominent peaks in each spectrum. These peaks are then organized into frequency trajectories by means of a peak continuation algorithm. The relevance of this algorithm is that it extracts the stable sinusoids present in the original sound (the deterministic component).

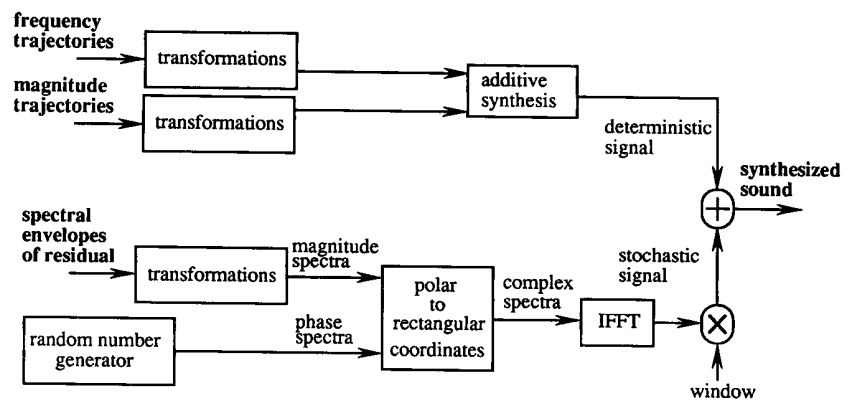


Fig. 3. This block diagram of the synthesis part of the Spectral Modeling Synthesis technique shows how to obtain a synthesized sound from the representation that resulted from the analysis process.

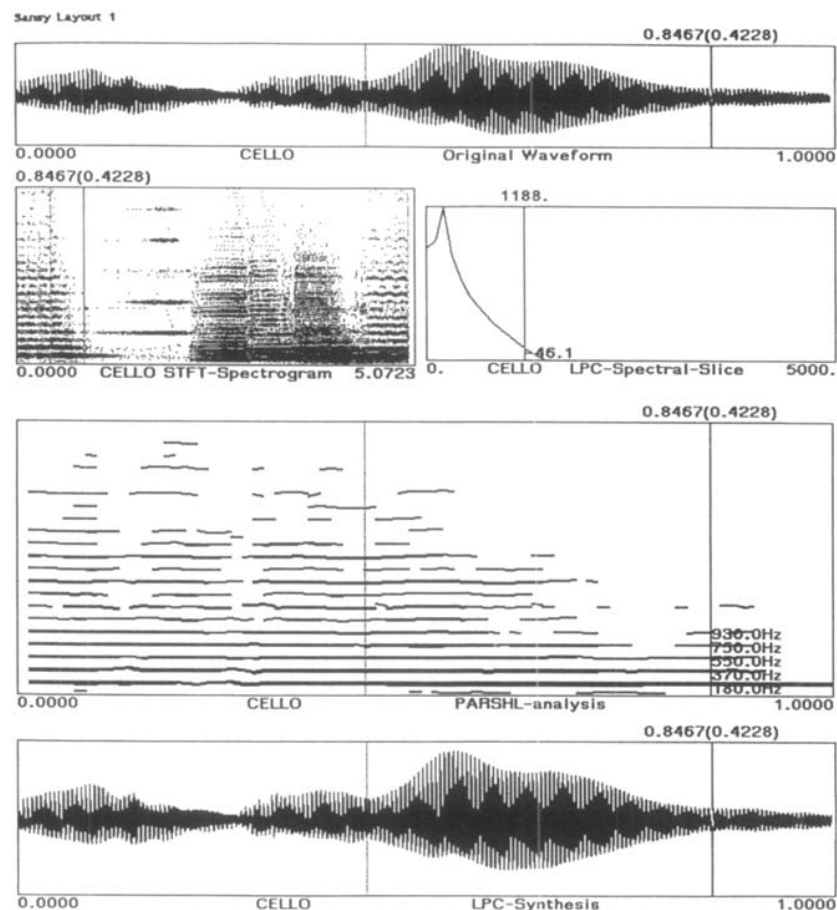


Fig. 4. This example of a typical display interface of the SANSY program shows an original cello sound at the top, three analysis representations in the middle and a synthesized version of the cello sound at the bottom.

The stochastic part of the waveform is calculated by first computing the STFT of the deterministic component, in the same way that the STFT of the original waveform was obtained, and then subtracting each magnitude spectrum from the corresponding spectrum of the original waveform. The envelope of each 'residual' spectrum is then derived by performing a line-segment approximation. These envelopes represent the stochastic signal.

The deterministic signal, i.e. the sinusoidal component, results from the magnitude and frequency trajectories, or their transformation, by generating a sine wave for each trajectory (i.e. additive synthesis).

The stochastic signal is the result of creating a complex spectrum (i.e. magnitude and phase spectra) for every spectral envelope of the residual, or its modification, and performing an inverse-STFT (using the overlap-add method to form the final output). The magnitude spectrum is the envelope itself, and the phase spectrum is generated by a random-number generator. This process corresponds to the filtering of white noise by a filter with a frequency response equal to the spectral envelope.

DESCRIPTION OF THE SANSY ENVIRONMENT

SANSY has been developed on a LISP-machine workstation, making use of an array processor for the signal-processing calculations [9]. A LISP machine is a computer built around the language LISP, a high-level computer language well suited for the specification and manipulation of abstract concepts. An array processor is a piece of hardware designed for mathematical computations. This combination of a LISP Machine with an array processor has the advantage of combining a highly sophisticated programming environment with a fast computation engine. The SANSY software is written in a dialect of LISP called ZETA-LISP, except for the low-level signal-processing routines, which are written in a low-level language using tools bor-

rowed from Speech and Phonetics Interactive Research Environment (SPIRE). SPIRE runs on a particular type of LISP Machine [10]. It is a program for interactively manipulating speech signals and computations on those signals. In addition, it can be used as a basis for developing other speech-processing systems and can be extended and customized by the user to perform specific tasks.

The SANSY environment includes a graphic interface that is completely configurable, from which the user controls analysis, transformation and synthesis processes. A typical SANSY-system display shows an original sound, several analysis representations and a synthesized version of the original sound (Fig. 4). The interface allows the user to modify data, including deterministic and stochastic representations, with graphic tools, as well as algorithmically. The possibilities for transformation are enormous—limited only by the imagination of the user. Apart from supporting the SMS technique, SANSY includes other standard tools for analyzing, editing and synthesizing musical sounds. SANSY is designed for use as a research environment rather than for musical composition. Musically it is used as a sound-modification program, although it is difficult to use it to combine or organize sounds in time. Therefore, once a sound has been created with SANSY, it has to be sent to a sound-editing and mixing program for integration into a musical composition.

The state of the SANSY system and of the SMS technique was, at the time that Austin composed *TT:TGE*, still under major development. Currently, the SANSY system is being rewritten so that it might be a useful tool for musicians, the major emphasis being composition. This new version will run on NeXT computers and be available to anyone interested.

References and Notes

1. M. Subotnick, "Silver Apples of the Moon", LP recording, Nonesuch Records H 71174.
2. N. Strauss, "Klaus Schoening: Radio You Can See", *Ear Magazine* 15, No. 3, 14–15 (1990).
3. R. Ashley, *Perfect Lives*, videotape recording, Lovely Music, New York, 1988. The keyboardist in

Perfect Lives is 'Blue' Gene Tyranny (aka Robert Sheff).

4. The premiere was broadcast live from a public performance at Hertz Hall, University of California, by radio station KPFA-FM, Berkeley, California, on 26 February 1990. It was performed by the University of California Chamber Chorus, Philip Brett, conductor; Chris Brown, keyboards; Chris Chafe, cello; and William Winant, Boie Radio Drum and percussion. The second performance was broadcast live from a public performance at the Merrill Ellis Intermedia Theater, University of North Texas (UNT), by radio station KNTU-FM, Denton, Texas, on 1 April 1990. It was performed by the UNT A Cappella Choir, Henry Gibbons, conductor; Carter Enyeart, cello; Steven Harlos, keyboards; and Robert Schietroma, KAT controller and percussion.

5. Spectral Modeling Synthesis was designed for use on the Symbolics LISP Computer.

6. Controller-instruments used in the first broadcasts included the Yamaha KX88 and the Roland A-80 Musical Interface Digital Interface (MIDI) digital keyboard controllers, the Boie Radio Drum (developed by Max Mathews at CCRMA) and the KAT digital percussion keyboard, and the cello (developed by Chris Chafe at CCRMA).

7. See Serra's technical explication starting on page 86.

8. See X. Serra, *A System for Sound Analysis/Transformation/Synthesis Based on a Deterministic plus Stochastic Decomposition*, doctoral thesis, later published as a technical paper by the Center for Computer Research in Music and Acoustics, Stanford University, 1989; X. Serra and J. O. Smith, "Spectral Modeling Synthesis: A Sound Analysis System Based on a Deterministic plus Stochastic Decomposition", *Computer Music Journal* 14, No. 4, 12–24 (1990).

9. SANSY is designed to run on the Symbolics LM-2, with the assistance of the FPS AP-120B array processor.

10. D. Shipman, "Development of Speech Research Software on the MIT LISP Machine", *The Journal of the Acoustic Society of America*, 103d meeting, Chicago, Illinois, U.S.A., 26–30 April 1982.

Bibliography

Austin, L., "Live-Electronic Music on the Third Coast", *Contemporary Music Review* 6, Part I (1990).

Austin, L., and T. Clark, *Learning to Compose: Modes, Materials and Models* (Dubuque, IA: Wm. C. Brown, 1989).

Austin, L., and E. DeLisa, "Modeling Processes of Musical Invention", *Proceedings of the 1987 International Computer Music Conference*, (San Francisco, CA: Computer Music Association, 1987) pp. 200–207.

Clark, T., "Coasts: On the Creative Edge with Composer Larry Austin", *Computer Music Journal* 13, No. 1, 21–35 (1989).

Clark, T., "Duality of Process and Drama in Larry Austin's Sonata Concertante", *Perspectives of New Music* 23, No. 1, 112–125 (1984).

Schwartz, E., and L. Starr, "David Tudor", *New Grove Dictionary of American Music* 4 (New York: MacMillan Press Ltd, 1986) pp. 418–419.

Strawn, J., ed., *Digital Audio Signal Processing: An Anthology* (Los Altos, CA: William Kaufmann, Inc., 1985).