CURRENT WORK AT CCRMA: AN OVERVIEW

1987?

Chris Chafe, Julius Smith, Patte Wood

Center for Computer Research in Music and Acoustics (CCRMA)
Department of Music, Stanford University
Stanford, California 94305

Introduction

The Stanford Center for Computer Research in Music and Acoustics (CCRMA) is an interdisciplinary facility and major resource where composers and researchers work together using the computer as a new musical and artistic medium, and as a research tool.

Areas of continuing research and development at CCRMA include: digital synthesis, signal processing, interactive composition, MIDI performance, graphics, machine recognition, psychoacoustics, and digital recording and editing.

The Center provides a specialized computer environment, fully networked for use of music synthesis and printing facilities accesible from a mainframe computer and several types of workstations, which include Symbolics and Xerox Lisp machines, 68000-based Unix workstations, and Apple MacIntoshes. The Center is capable of serving a number of composers and researchers simultaneously, providing for each a direct control over the musical medium to a degree which was never before possible. In addition to the Stanford community, CCRMA serves as a resource for visiting composers and researchers from throughout the world. The results of this interaction are presented to the public through concerts, demonstrations and research publications.

A summary of the current major work effort at CCRMA follows.

Synthesis Algorithms: Existing Techniques

A number of sounds synthesis algorithms in daily use at CCRMA are increasingly available in commercial synthesis gear. For example, frequency modulation is found in the Yamaha "X" line of synthesizers. Other synthesizers are being engineered for additive (Fourier) synthesis, waveshaping, sampling, the "Digitar," and "Chant" (formant synthesis). Research on these techniques at CCRMA continues to extend their "naturalness" and ease of manipulation, and to create better computer models of acoustic instruments. Many automatic aids are now available to simplify the analysis/ data reduction/ synthesis chain common in model-building work.

A compiler is being written to allow porting of Yamaha DX-7 voicing data to CCRMA's general purpose synthesizer (the "SAMSON BOX"). With this facility, DX voices will be available for study under varying conditions: alternative pitch scales, different sampling rates, and other styles of control envelopes. It will provide composers at our studio with the ability to create custom voices on a DX-7 and have those voices available for manipulation by a music composition or general-purpose programming language. (S. Krupowicz)

A dynamic-spectrum additive synthesis method is in use. Fundamental timbres are derived from a database of analyzed instrumental and vocal steady-state spectra. An additive synthesis instrument enables arbitrary continuous interpolation between arbitrary mixes of these spectra. In addition, various ways of modulating the spectra or spectral mixes are provided. (M. McNabb)

An analysis method, PARSHL, follows the peaks over time in a series of spectral frames. Each frame consists of magnitude data from the FFT of a windowed sampled data segment. Additive synthesis is used to recreate the measured amplitude and frequency trajectories. Simulation of non-harmonic tones is very good. (J. Smith)

New Synthesis Algorithms: Physical Models

Most synthesis techniques provide the computer musician with sound that is controlled via *acoustic components* (such as partials). Patches or computer "instruments" are configurations of generators of these components, eg. sine oscillator. A new approach models real instruments directly, using a mechanical simulation to produce the complex sound waveform. The "handles" on such a computer instrument are those which are already familiar to a player: bow velocity, embouchure, etc.

There exists a single mathematical model for efficiently simulating all reed instruments, bowed strings, brasses, and possibly switching air-jet instruments such as the pipe organ and flute. The model consists of one or more "negative-resistance" nonlinear oscillators coupled bidirectionally to a linear resonator. The nonlinear oscillators act as reeds, bow-hairs, lips, or air-jets. The linear filter functions as a bore, body, horn, or pipe, together with the effects of fingering.

Signal modeling work in music is related to techniques with application to speech, seismic disturbance, light propagation, thermal gradient, structural stress, aircraft motion, stock market level, weather patterns, or the sensation in a robot's hand. A model for the violin, based on underlying physics has been developed using new techniques for digital filter design and system identification. Three important components of the violin deserve individual consideration: (1) The violin body contributes to the timbre of the sound by shaping the spectrum in a fixed way. (2) The string of the violin, well-modeled as a linear filter, but standard filter types are far too expensive to consider in this context. (3) The bow-string interaction, a nonlinear phenomenon. (J. Smith and D. Jaffe)

This model performs best when controlled in a player-like fashion. To simulate a player's interpretation, musical scores have been converted into sequences of tablature-like motions and phrasing controls. The software is essentially an "expert-system" that creates time-varying envelopes for the instrument controls. (C. Chafe)

New Synthesis Algorithms: Hybrid Methods

A model for piano tones is being developed that incorporates the most advantageous aspects of several synthesis techniques including additive synthesis, frequency modulation synthesis, and filtering in an effort to maintain both economy and accuracy throughout the range of the piano. An effort has been to provide a high degree of flexibility and control by incorporating a "building block" type of synthesis whereby the user can readily incorporate either or both of the synthesis techniques at any location within the spectral range. In addition, careful attention has been paid to the stretching of the harmonics both in additive synthesis, and within spectral groups modeled using FM synthesis. (J. Polito and J. Marks)

Another hybrid approach to modeling has been taken to simulate bar percussion instruments. The starting point is current knowledge of the acoustics of the instruments. Then a modular analysis/synthesis process is taken based on Discrete Fourier Transform techniques and filter methods. The result of the research is an algorithm implemented on the Systems Concepts Digital Synthesizer ("Samson Box") that integrates several synthesis strategies. This algorithm maps to the underlying physics of the instruments and its parameters can be easily specified. (X. Serra)

Signal Processing

Interest in signal processing has led to improvements in digital filter design for musical applications. Another area has been development of better tools for analysis of musical signals. The center's library of general signal processing routines available online at the center is probably unsurpassed in terms of completeness.

A digital resampling method has been developed which allows non-uniform and time-varying resampling. The method is based on interpolated look-up in a large table of filter coefficients. One filter table handles all conversion factors. (J. Smith)

Digital filters are derived by sampling lossless propagation through a network of ideal "waveguides." Sampled waveguide networks provide numerically robust time-varying filter structures which we call "WaveGuide Filters" (WGF). (J. Smith)

One desirable property of waveguide filters is that signal power can be decoupled from changes in the filter parameters. The WGF structures can be "balanced" such that the decoupling between signal power and time-varying filter coefficients is maintained for each individual section in the structure. (A "section" is a stretch of ideal waveguide material.) (J. Smith)

Numerical round-off effects are easily controlled in WGF structures. For example, it is simple to suppress limit cycles and overflow oscillations, even in the time-varying case. This is accomplished by using "passive" arithmetic; the exact physical interpretation of signals in a WGF make it clear how to define passive arithmetic. (J. Smith)

Finally, WGF structures can be interconnected in cascade or in parallel without disturbing the signal/coefficient decoupling, signal power balance, or roundoff passivity properties. Thus, waveguide filters are very useful for modeling physical systems, and the exactness of their physical interpretation enhances their suitability for the time-varying case. It is interesting that these more elaborate physical modeling networks do not reduce to standard filter structures. All results are obtained for the multi-input/multi-output case. (J. Smith)

A new type of digital reverberator is being developed which is based on closed networks of intersecting waveguides. Each waveguide is a bi-directional delay line, of arbitrary length, and each intersection (of any number of waveguides) produces lossless signal scattering. By creating a closed network of waveguides, the total signal energy in the structure is preserved. A reverberator is constructed by introducing small losses in the network to achieve a desired reverberation time. The inputs and outputs can be chosen anywhere in the structure. There are many reasons to construct reverberators (or any recursive digital filter for that matter) from lossless waveguide networks: (1) the scattering junctions can be made time varying without altering stored energy, (2) an "erector set" for lossless networks is obtained, allowing any number of branches to be fitted together in any desired configuration (with changes allowed in real time), (3) limit cycles and overflow oscillations are easily eliminated, regardless of interconnection, (4) an exact physical interpretation exists for all signals in the structure, and (5) the implementation is computationally efficient. (J. Smith)

A convenient structure has been determined for implementing digital "phasers," "flangers," "comb filters," and the like. These sound modifiers all work by sweeping "notches" through the spectrum of a sound. The main feature of the structure is that a fixed number of notches is provided which can be controlled independently. Each notch-section is made using a second-order allpass filter. (J. Smith)

While digital filter design is a relatively mature field, there is a wide gap between generally available design criteria and design criteria most appropriate in audio applications. Three to match more closely the characteristics of hearing: (1) The desired frequency response is "warped" so that critical bands of hearing have bandwidths largely independent of center-frequency; (2) the desired frequency response is smoothed and weighted according to hearing sensitivity; and (3) the digital filter is fit on a dB scale instead of a linear one. Each of these steps is based on approximations, but each can significantly improve the quality of the resulting filter design. (J. Smith)

Research is in progress to build a programming environment for signal processing development. It consists of a signal-processing library, a natural mathematical front-end syntax, and a comprehensive "HELP" facility. This will result in an interactive signal-processing "work bench." The purpose of the facility is to provide convenient interactive access to the CCRMA signal processing library and to provide an expandable programming environment for signal processing applications. (J. Smith)

In analyzing musical signals, the most critical signal processing step is the transformation of the sampled input signal to a frequency-domain representation. The Fast Fourier Transform is attractive for its efficiency, but the linear (as opposed to exponential) spacing of its frequency bins yields poor discrimination in the lower frequency range, compared with the ear's. In order to improve this situation while retaining the FFT's speed, we have devised a method of time-varying spectral analysis, called *Bounded-Q Frequency Transform* (or BQFT) which uses FFTs within a recursive scheme of octave decimation. The BQFT offers better than semitone resolution, which is sufficient for the effective processing of polyphonic textures by the further levels of analysis. (B. Mont-Reynaud)

Composition: Music Programming Languages

The computer provides a tool for composition as well as sound synthesis. Pieces created in the studio at CCRMA commonly invoke computer aids at some level, to help control the details of the chosen musical materials. A standard programming language has evolved for this work and has been used by a large number of composers and acousticians. Also listed below are several topics of musical research of common interest to computer musicians.

We have taken several steps toward providing composers with the computational resources they need to write computer music. The first step was the language Pla. Pla's primary goal was to give composers a music programming language powerful enough to express any compositional algorithm in terms congenial to the individual composers. Pla is a programming language in the Algol family. Besides the Algol part of SAIL, Pla also includes implementation of lists, contexts, flavors, and several musically useful constructs. Our experience with Pla has demonstrated that many musical notions can be easily and compactly expressed by a standard computer programming language. (W. Schottstaedt)

Once the algorithms could be expressed, fancier editing and display facitilities were needed. These were provided by Edpla which could be described as a graphics-oriented music editor based on Pla. Pla and Edpla put increasingly complicated musical situations within reach, but debugging these situations began to pose major problems. The third step, therefore, was the development of a group of debuggers (EdSam, Delila, Debugh). At this point, composers began to complain about the muddle of languages and mutually inconsistent conventions the various programs used. Although EdPla and EdSam were designed to be as close as possible to E (our main text editor), and Pla was meant to be very close to SAIL (our main programming language), there were a number of ugly conventions left over from previous times. Also some ideas that seemed good in prospect turned out to be less so in retrospect. Our next step, therefore, was an attempt to build a new compositional world based on one language (SAIL) and a small set of consistent metaphors. This was Jetsam, an object oriented approach to computer instrument design and note list handling. After several years, however, we have reached the point where simple additions to Pla and Jetsam are no longer sufficient. To go further requires a substantially more powerful software environment than is available via our mainframe system. LISP workstations are now being used toward this goal. NED is a multiple representation note-list editor implemented in this new environment. (W. Schottstaedt and D. Jaffe)

Computer music systems have tended to deal in a clumsy manner with both inter- and intra-voice timing. A straightforward solution called a "time-map" is under way. It provides a general formalism for describing ensemble timing interaction and allows for both arbitrarily free tempo variations and precisely specified simultaneities. The time map method has proven useful both for the simulation of expressive playing and for compositional timing manipulation. (D. Jaffe)

Two programs were written to aid composers interested in microtonal scales and nonstandard tuning systems. The first retunes the frequencies in a PLA note list to any given system. The second, which extends earlier programs for real-time control of the Samson Box, uses the terminal keyboard as a musical keyboard with user-programmable tuning and key layout, and graphic display of the pitches assigned to each key. Various routines including a sequencer generating PLA code are provided. (D. Keislar)

Fractal geometry is a recently-developed area of mathematics. It has been found to have a close relationship with many natural phenomena which were difficult to model with traditional mathematics. Many musical applications of fractals are being experimented with, including the generation of melodies, rhythms, dynamics, vibrato, spectra, and spectral modulation. (M. McNabb)

Species counterpoint as presented by J. J. Fux in "Gradus Ad Parnassum" appears to be a ready made case for a rule based "expert system". In programs of this sort, knowledge is encoded as a list of IF..THEN statements. Using these statements, a computer program can find acceptable solutions to species counterpoint problems. A program has been developed that can write species counterpoint in any of the standard modes and with up to six voices. (W. Schottstaedt)

Composition: Works and Activities

In general there were more then thirty pieces realized at CCRMA during the academic year 1985-1986 by both resident composers and guest composers. Pieces were presented in major festivals both in the USA and elsewhere: Rick Taube's JubJub was performed at the Tanglewood Music Festival, Amnon Wolman's M for chamber orchestra and tape, was read at the Aspen Music Festival. A concert of CCRMA's music which included: David Jaffe's Silicon Valley Breakdown and Bristlecone Concerto #3 for 10 players and tape, Bill Schottstaedt's Colony V and Doug Fulton's Red Cup and Rat was presented at the Cabrillo Music Festival. Guest Composers Ira Mowitz piece Jublum was selected be the ISCM American Jury to represent the USA in the world ISCM Music days at Koln Germany. Other perfomances of pieces created at CCRMA included the GRM/INA series in Paris and a concert sponsored by the Studio 200 in Tokyo.

In December 1985 a premier of guest composer Mike McNabb's work Invisible Cities was presented at Stanford. Supported by an NEA grant this piece for piano, saxophone, computer tape, and electronics, is a forty minutes ballet for robots and dancers created for the ODC/San Francisco dance company. Composers Bill Schottstaedt, Chris Chafe and David Jaffe were all represented at the Dedication concert of the new CCRMA facilities in June 1986. The concert included John Chowning's Turenas, Chris Chafe's In A Word for 'cello and tape, Bill Scottstaedt's Water Music, David Jaffe's Impossible Animals for choir and tape (commissioned by the Hamilton College Choir), and Jonathan Bergers Diptych, for mezzo soprano, string quartet and tape (performed by the prestigious Stanford String Quartet). In July 1986, guest composer Johannes Goebel presented the U.S. premier of his piece Passage, for percussionist, speakers, and electronics, in a concert co-sponsored by the Goethe Institute of San Francisco.

The annual CCRMA computer music concert, at Frost Amphitheater, was held in July 1986. This concert featured guest composer Jean Claude Risset's Sud, other pieces on the concert included Richard Karpen's Eclipse, Adolfo Nunez's Images for trumpet, trombone, horn, piano, soprano and computer-generated tape (winner of the 2nd prize at the Hanna Prize composition competition), Matthew Fields' Walk and Amnon Wolman's A Circle in the Fire, for bass clarinet and tape.

MIDI

With the rapid development of increasingly sophisticated synthesis-processing hardware in the music industry, the MIDI interconnection protocol becomes an important basis for building systems. Small often inexpensive digital devices, (synthesizers, reverberators, signal processors, etc.) can be assembled in a modular form to arbitrary depths of complexity through the common MIDI control protocol. Advantages of such systems over monolithic systems are compactness and reliability resulting from VLSI implementation of the hardware and nearly universal availability. Only through such systems can real-time performance of computer music be effectively realized. The disadvantage of these systems lies in their lack of generality which will only find a remedy through the continued development of the semi-condutor technology, independent of our influence, and the music industries understanding as to where available power is best applied, certainly not independent of our influence. CCRMA is, therefore, developing a MIDI-based studio in order to exploit the existing technology and participate in the definition of future technology based upon long experience in synthesis, signal-processing, and the study of important links between the physical and perceptual attributes of sound. (J. Chowning)

Computer control of MIDI data often deals excusively with keyboard-style musical events. The MIDI format itself is not limited in this way. Of interest are methods that would represent performance gestures on string and wind intruments. A system is under development for capture and control of violin bow gestures. Expressive qualities of bow speed and pressure will be tied to MIDI "aftertouch" parameters. (C. Chafe)

Music Graphics

High resolution music printing at CCRMA is linked with an interactive score editing language. Notes are input from the typewriter keyboard (or from the output of a computer program such as the acoustic transcriber) and displayed on a music stave. As with a text editor, data is readily polished into shape and certain automatic formatting aids are available. Justification and beaming of rhythms, page layout and part extraction are examples. Printed output is engraver quality. The system is being ported to personal computer systems. (L. Smith)

"Visible Music" refers to a system that displays high resolution spectograms and allows the user to interactively study features found in the display. A cursor reveals frequency and time information and can be used to collect notes from the display, updating a notelist when clicked. A harmonic cursor is used to highlight harmonically related groups of spectral lines. The tool is valuable for hand analysis of acoustic data. (B. Mont-Reynaud)

Music Analysis and Machine Recognition

Experiments in automatic music recognition at CCRMA have been in progress for five years. Digitized sound recordings of instrumental music are analyzed and transcribed by computer using artificial intelligence techniques. The current effort is directed at polyphonic examples with a variety of instruments and musical styles. The overall goal of the work is to provide a tool for the study of musical performance, for applications requiring tracking of live musicians, for manuscript work and for segmentation of digital audio recordings.

In transcribing performed music, a system must extract every note played, identifying timings, pitch, dynamic information and other parameters. A capability for source discrimination is also required if a polyphonic musical texture is present, that is, if the signal results from multiple simultaneous sources. Identification algorithms which are adequate for monophonic input are not as effective and the balance between underdetection and overdetection is more difficult to achieve. Improved performance is possible by combining the strengths of alternative algorithms for each identification task. The use of acoustic knowledge and context generated at higher levels also assists in the recognition process. As signal history accumulates, a particular event need not be as visible for successful identification if it clearly fits into an already established context. (C. Chafe and D. Jaffe)

Attempts are being made to capture some aspects of the perception of musical patterns in formal definitions and usable algorithms. Here, attention is restricted to a single musical line. As an example, viewing a piece as a string of features, patterns are defined as substrings that recur. Algorithms for the efficient enumeration of patterns, and for the selection of preferred patterns have been created. For example, ABRACADABRA yields 9 patterns. The preferred patterns are ABRA and A. The other 7 patterns are implicit in ABRA. The arrangement of patterns instances provides strong clues for the metrical organization of a piece. (B. Mont-Reynaud)

Psychoacoustics, Ear Modeling, Data Compression

An analysis/synthesis system allows interpolation between familiar, naturally-occuring tones. In exploring musical timbre space, parameters under control are the overall amplitude envelope, the pitch function and the short-time spectral content of the tone. Automated data reduction is applied to the analyzed data. It is hoped that the tool will be useful for controlled generation of new timbres of musical interest. (D. Lo)

Another study utilizes a time-domain analysis of speech signals for both recognition and compression. Analysis is based on perceptually relevant features of the waveform. Low bit rates are achieved, and the reconstructed signal preserves the quality of the individual speaker. The techniques are also valuable for analysis of musical sound. Short-term events that are analogous to speech consonants can be effectively analyzed, such as starting transients of winds or strings. (B. Shannon)