Center for Computer Research in Music and Acoustics (CCRMA)
The Knoll, Stanford University
Stanford, California 94305-8180
Phone: (415) 723-4971 Fax: (415) 723-8468

OVERVIEW

Center for Computer Research in Music and Acoustics

APRIL 1994

edited by

Patte Wood and Scott Van Duyne

Research sponsored in part by

The Mellon Foundation
The National Endowment for the Arts
The Rockefeller Foundation

CENTER FOR COMPUTER RESEARCH IN MUSIC AND ACOUSTICS
Department of Music, Stanford University

OVERVIEW
May 1994

TABLE OF CONTENTS

CCRMA ROSTER

PPN	Name		
Staff & Faculty			
JC	John Chowning, Professor of Music, Director		
CC	Chris Chafe, Technical Coordinator/Assoc. Prof. of Music		
PATTE	Patte Wood, Administrative Director		
HMK	Heidi Kugler, Secretary		
NANDO	Fernando Lpez Lezcano, Scientific Research Programmer		
JAY	Jay Kadis, Audio Engineer		
BAU	Marcia Bauman, Research Associate		
PRC	Perry Cook, Research Associate		
MVM	Max Mathews, Professor of Music (Research)		
JRP	John Pierce, Visiting Professor of Music, Emeritus		
BIL	William Schottstaedt, Research Associate		
EDS	Earl Schubert, Professor of Speech and Hearing, Emeritus		
LCS	Leland Smith, Professor of Music		
JOS	Julius Smith, Assoc. Prof. of Music		
Student Research Assistants		Degree	Project
DPBERNER	David Berners	PhD EE	Research
BRENT	Brent Gillespie	PhD ME	Research
SCOTTL	Scott Levine	PhD EE	Research
PUTNAM	William Putnam	PhD EE	Research
AVERY	Avery Wang	PhD EE	Research
LZ	Li Zeng	DMA Music	Composition
Student Work Study		Degree	Project
CURTIS	Ethan Eldridge	AB MST	Work Study
MEJANE	Jane Rivera	AB MST	Work Study
OTTO	Chris Otto	AB MST	Work Study
WHITLOCK	Rob Whitlock	AB MST	Work Study
Graduate Students		Degree	Project
AGUIAR	Celso Aguiar	DMA Music	Composition
YORGOS	Yorgos Antoniou	DMA Music	Composition
JAN	Jan Chomyszyn	PhD Music	Research
KUI	Kui Dong	DMA Music	Composition
JMD	Janet Dunbar	DMA Music	Composition
MICHAEL	Michael Edwards	MA,DMA Music	Composition
RJFLECK	Robert Fleck	PhD Music	Research
NICKY	Nicholas Hind	DMA Music	Composition
HOPKINS	Nicholas Hopkins	DMA Music	Composition
ALEDIN	Alex Igoudin	PhD Music	Research
ANIA	Anna Marchwinska	DMA Music (Piano)	
EIG	Enrique Moreno	PhD Music	Research
NORTON	Jonathan Norton	DMA Music Theory	Research
GARY	Gary Scavone	PhD Music	Research
DBS	David Soley	DMA Music	Composition
ATAU	Atau Tanaka	PhD Music	Research

WISHBON	Steven Trautmann	DMAMusic	Composition
RVK	Rick VanderKam	PhD EE	Research
JANVAN	Jan VandenHeede	DMA Music	
RVK	Rick VanderKam	PhD EE	Research
SAVD	Scott Van Duyne	PhD Music	Research
PVH	Paul von Hippel	PhD Music	Research
DDZ	David Zicarelli	PhD Hearing and Speech	Research

Undergraduate & MST Majors		Degree
COULOUR	Kwame Anku	MST
	Jeff Chen	MST
SADMAN	Stephen Davis	MST/EE
PHRH	Peter Habicht	Acoustical Engineering
	William Hogan	MST
CMGOODAN	Charles Goodan	MST
MONTERAY	Monte Klaudt	MST
MEISEL	Ben Meisel	MST
BOOMRANG	Jason Naiden	MST Indtustrial Engineering
SAWYER	Jim Sawyer	MST Math Computational Science
	Seth Spalding	MST
ODS	Owen Smith	MST
GRSMITH	Greg Smith	CS
TUCKER	Blake Tucker	MST
TRACE	Trace Wax	Music
WITCHEL	Emmett Witchel	MST

Visiting Scholars/Composers - 1993-94

RUBEN	Ruben Altschuler	Visiting Composer, Argentina
EITAN	Eitan Avitsur	Visiting Composer, Israel
EMB	Emily Bezar	Independent composer and performer
MAB	Marina Bosi	Visiting Researcher,Dolby Labs
BREGMAN	Albert Bregman	Visiting Researcher, McGill University
JDC	Joanne Carey	Visiting Composer
NPCARTER	Nicholas Carter	Visiting Researcher, University of Sheffield
DESAIN	Peter Desain	Visiting Researcher, Holland
RD	Richard Duda	Researcher, San Jose State University
JEG	Johannes Goebel	Visiting Composer, ZKM, Karlsruhe, Germany
HONING	Henkjan Honing	Visiting Researcher, Holland
DAJ	David Jaffe	Visiting Composer
DK	Doug Keislar	Visiting Researcher
BK	Ben Knapp	Visiting Research, San Jose State University
KAZUKI	Kazuki Kuriyama	Composer, Kunitachi College Japan
PEER	Peer Landa	Composer, Norway
LEVITIN	Dan Levitin	Visiting Researcher, University of Oregon
LEWISPSC	Philip Lewis	Independent researcher
LUKAS	Lukas Ligeti	Visiting composer, Austria
HSL	Hugh Lusted	Visiting researcher, BioMuse
FLM	Fred Malouf	Independent composer
JRM	Janis Mattox	Independent composer
DMILLEN	Dale Millen	Visiting Computer/Ressearcher, U. of Arkansas
BMR	Bernard Mont-Reynaud	Visiting researcher
MULLER	Carl Muller	Visiting researcher

SILE	Sile O'Mondhrain	Visiting composer/research, Ireland
TO	Tom Oberheim	Visiting researcher
JUAN	Juan Pampin	Visiting composer, Argentina
STP	Stephen Pope	Visiting composer/editor Computer Music Journal
NICK	Nick Porcaro	Independent researcher
ROC	Davide Rochesso	Visiting researcher, Italy
XJS	Xavier Serra	Visiting researcher, Spain
STEPHEN	Stephen Schwanauer	Composer
NS	Naoki Shibata	Visiting researcher, NEC, Japan
MALCOLM	Malcolm Slaney	Visiting researcher, Apple Computer
WENYUSU	Wen-Yu Su	Visiting researcher, ITRI, Taiwan
HKT	Heinrich Taube	Visiting composer
MARCO	Marco Trevisani	Visiting composer, Italy
AYAKO	Masato Yako	Visiting research, Kyushi Institute of Design, Japan
DAW	David Wessel	Professor of Music, UC Berkeley

CCRMA

The Stanford Center for Computer Research in Music and Acoustics (CCRMA) is an interdisciplinary facility where composers and researchers work together using computer-based technology both as a musical and artistic medium and as a research tool.

Areas of ongoing interest at CCRMA include: Applications Hardware, Applications Software, Synthesis Techniques and Algorithms, Physical Modeling, Real-Time Controllers, Signal Processing, Digital Recording and Editing, Psychoacoustics and Musical Acoustics, Music Manuscripting by Computer, Composition, and Real-Time Applications.

The CCRMA community consists of administrative and technical staff, faculty, research associates, graduate research assistants, graduate and undergraduate students, visiting scholars and composers, and industrial associates. Major departments actively represented at CCRMA include Music, Electrical Engineering, Mechanical Engineering, Computer Science, and Psychology.

Center activities include academic courses, seminars, small interest group meetings, summer workshops and presentations. Concerts of computer music are presented several times each year with an annual outdoor computer music festival in July. In-house technical reports and recordings are available, and public demonstrations of ongoing work at CCRMA are held periodically.

Research results are published and presented at professional meetings, international conferences and in established journals including the Computer Music Journal, Journal of the Audio Engineering Society, and the Journal of the Acoustical Society of America. Compositions are presented in new music festivals and radio broadcasts throughout the world and have been recorded on cassette, LP, and compact disk.

Support for CCRMA has been received from the late Doreen B. Townsend, the California Arts Council, the Ann and Gordon Getty Foundation, the Mellon Foundation, the National Endowment for the Arts, the National Science Foundation, the Rockefeller Foundation (for artists-in-residence), the System Development Foundation, Apple Computer, Bio Control, Crystal Semiconductor, E-mu, Dynacord, ITRI CCL Taiwan, Korg, Matsushita, Media Vision, NEC, NeXT Inc., Opcode, Rockwell Int'l, Roland, Symbolics, Xerox Palo Alto Research Center, Yamaha, Zeta Music Partners, and private gifts.

Education

CCRMA is an interdisciplinary center administered by the Music Department at Stanford University. Classes and seminars taught at the center are open to registered Stanford students and visiting scholars. The facility is also available to registered Stanford students and visiting scholars for research projects which coincide with ongoing work at the center.

Prospective graduate students especially interested in the work at CCRMA should apply to the degree program at Stanford most closely aligned with their specific field of interest, i.e. Music, Computer Science, Electrical Engineering, Psychology, etc. Graduate degree programs offered in music are the DMA in Composition, and the PhD in Computer-Based Music Theory and Acoustics. Acceptance in music theory or composition is largely based upon musical criteria, not knowledge of computing. Admission requirements for degree programs can be obtained directly from each particular department. CCRMA does not itself offer a degree.

The Music department offers an undergraduate major in Music, Science, and Technology. The major is designed for those students with a strong interest in the acoustic and psychoacoustic foundations of music and the musical ramifications of the rapidly evolving computer technology. The program is highly multi-disciplinary and entails a substantial research project under faculty guidance. This program can serve as a complementary major to students in the sciences and engineering. Requirements for the undergraduate major are available from the Stanford Music Department.

Courses offered at CCRMA include:

- Introduction to Computer Music;
- Introduction to Composition and Programming Using MIDI-Based Systems;
- Computer-Generated Music – (topics include Fundamentals of Computer-Generated Sound, Compositional Algorithms, Psychoacoustics, and Spatial Processing);
- Seminar in Signal Processing for Music Research – (topics include The Discrete Fourier Transform, Applications of the Fast Fourier Transform, and Physical Modeling of Musical Instruments);
- Psychophysics and Cognitive Psychology for Musicians – (topics include the basic concepts and experiments in psychophysics and cognitive psychology which are relevant to the use of sound, and especially of synthesized sound in music);
- Theory and Practice of Recording – (topics include elementary electronic, physics of transduction and magnetic recording of sound, acoustic measurement techniques, operation and maintenance of recording equipment, and recording engineering principles);
- Seminars on Sound Perception and Special Topics.

CCRMA also offers a series of summer workshops. Courses offered are:

- Introduction to Yamaha VL Synthesizers and Algorithmic Composition in Common LISP;
- Advanced Projects in Algorithmic Composition;
- Intensive Music/Audio Digital Signal Processing and Spectral Modeling Synthesis;
- Music Printing with Small Computers using SCORE.

Facilities

CCRMA located on the Stanford University campus in a building that was refurbished in 1986 to meet its unique needs. The facility includes a large quadraphonic experimental space with adjoining control room/studio, an all-digital recording studio with adjoining control room, a MIDI-based small systems studio, several work areas with workstations, synthesizers and speakers, a seminar room, an in-house reference library, classrooms and offices. The building has been wired so that any office or workspace can connect with the workstation network. A gateway connects the workstations to the campus at large and also to national and international networks. A description of the hardware and software environment follows below.

The CCRMA hardware environment consists of an Ethernet network of workstations running the NextStep operating system (both NeXTs and Intel based PCs) and Macintosh computers, and a LocalTalk network for some of the Macintoshes. Both networks are interconnected by a Shiva FastPath box. Digital signal processing is possible on the NeXT computers, both via built-in Motorola 56001 DSP hardware and on three Ariel Quint Processor boards which contribute five additional 56001 processors each. The Macintosh systems also provide DSP via Digidesign Sound Accelerator boards. MIDI input and output are supported from Macintosh and NeXT computers. Digital audio processors include a Studer-Editech Dyaxis II system which can convert all popular digital audio formats as well as store and edit audio digitally, a Sony PCM-1610 3/4" video PCM system for mastering Compact Disks, several Singular Solutions analog and digital audio input systems for the NeXTs, and several Panasonic and Sony DAT recorders. Text and graphics are handled by an HSD color scanner on the NeXT system and by both a NeXT Laser printer and an Apple Laserwriter.

The MIDI-based systems include various Macintosh computers with Yamaha, Roland and Korg equipment including Yamaha DX, TX, SY, TG and VL synthesizers, KX88 keyboard controller, Korg WaveStations, an E-mu Systems Emulator III sampler, and digital delays and reverberation. Other equipment available includes IVL pitch trackers, a Buchla Lightning MIDI controller, a Mathews Radio Drum controller, a JL Cooper MIDI patcher, Akai 32 channel MIDI-controlled audio patchbay and drum machines from Yamaha and Roland.

Studio recording equipment includes a 24 track mixer, an 8 track TEAC analog recorder, a Yamaha DMR8 digital recorder and mixing console, a TEAC 8 track digital recorder, various signal processing devices, Westlake monitor speakers and an assortment of high quality microphones.

The CCRMA software has been developed over a twenty-year period, and consists of a vast set of programs and system tools for editing, viewing, synthesizing, and analyzing sound. Much of the software was originally written in SAIL, a sophisticated Algol-like language for use on the previous mainframe and has been ported to the new workstation environment. Recent projects in music recognition, audio, signal processing, acoustics, psychoacoustics and physical modeling have been developed in languages native to the workstations, primarily Common Lisp, Objective-C, and Smalltalk.

RESEARCH

Introduction

Computer music is an interdisciplinary field. The research summarized in this overview spans such areas as engineering, physics, computer science, psychology, and music (including performance, analysis, and composition); and, any given specific research topic may require sophistication in several of these fields. This document can only contain a brief review of the work beeing done at CCRMA. For a more complete description of the research, a list of CCRMA research publications is included. Copies of reports are available upon request.

Although the research described in this document is diverse, the following rough grouping of the work presented here may be useful: Music composition software development continues (Schottstaedt, Taube, Lopez, Oppenheim), and new systems have been created for rapid prototyping of signal processing and sound systhesis algorithms (Smith, Jaffe, Porcaro). Research in physical modeling of musical instruments and reveberant environments continues to grow (Smith, Fleck, Rocchesso, Van Duyne, Cook, Chafe, Berners, Scavone, Pierce, Su). Innovative signal processing approaches are being studied (Smith, Wang, Levine). Much attention is currently being given to the problem of human control of computers and sound synthesis algorithms running on computers (Chafe, Cook, Morrill, Stilson, O'Modhrain, Gillespie, Mathews, Lopez, Jaffe, Tanaka, Putnam). In the psychoacoustics area, work continues in sound perception research and its applications (Trautmann, Chomyszyn, Pierce, Levitin, Shepard, Cook, Bregman). In addition, CCRMA now houses two archival research libraries in electro-acoutic music and in acoustics (Mathews, Bauman, Scavone).

The researchers working at CCRMA include graduate students, faculty, staff, and visiting scholars. Email may be sent to any of these people by mailing to *login name*@CCRMA.STANFORD.EDU, where *login names* are listed in the roster elsewhere in this publication.

COMPOSITIONAL ACTIVITIES

Since the late 60's most of the work in composition at CCRMA has been done in a software environment which evolved from the Music V program originally developed at Bell Labs by Max Mathews and his research group. The hardware and software has improved over the decades following, and the names of things have changed. Ported to a PDP10, Music V became the Mus10 music compiler system and played scores composed in Leland Smith's SCORE language. The compiler was replaced in 1977 with dedicated synthesis hardware in the form of the Systems Concepts Digital Synthesizer (built by Peter Samson and known as "the Samson Box"). The Samson Box was capable of utilizing many types of synthesis techniques such as additive synthesis, frequency modulation, digital filtering and some analysis-based synthesis methods. The PLA language, written by Bill Schottstaedt, allowed composers to specify parametric data for the Samson Box as well as for other sound processing procedures on the PDP10 mainframe (and its eventual replacement, a Foonly F4). On April 3, 1992, the Foonly and Samsonbox were officially retired. CCRMA has transitioned to a network of workstations (NeXT and Intel based PC's) running the NextStep operating system and the functionality of PLA exists now in the form of Common Music and Stella (written in Common Lisp by Rick Taube), a software package that can write scores by listing parameters and their values, or by creating algorithms which then automatically determine any number of the parameters' values. Common Music (CM) can write scores in several different syntaxes (currently CLM, CMN, MusicKit, MIDI, CSound and

Paul Lansky's real-time mixing program RT). The scores can then be rendered on workstations using any of the target synthesis programs. For example, CLM (Common Lisp Music, written by Bill Schottstaedt) is a widely used and fast software synthesis and signal processing package that can make use of multiple Motorola 56001 DSP's. CCRMA has also become the maintainer and distributor of NeXT's MusicKit, a real time toolkit for computers running NextStep that merges the MIDI and real time synthesis paradigms and can also be the target of Common Music generated scores.

Continuity has been maintained over the entire era. For example, scores created on the PDP10 or Samson Box have been recomputed in the NeXTStep computing environment, taking advantage of its increased audio precision. To summarize all these names for CCRMA's composing environment, the synthesis instrument languages have been called Mus10 - Sambox - CLM / MusicKit and the composing language succession has been SCORE - PLA - Common Music / Stella. Other computers and software are also used for composition. Several composers have realized pieces which make extensive use of MIDI equipment, especially Yamaha synthesizers controlled via Macintosh computers. The acquisition of a Dyaxis II Digital Audio Processor and several Macintosh II computers has brought renewed interest in real-time control and computer-based "musique concrète." The programming environments being used for composition and developmental research for all these systems include MAX, LeLisp, Smalltalk, Common Lisp, DMIX (a flexible compositional environment, written by Dan Oppenheim), Objective C and C.

Since its beginning, works composed at CCRMA have been highlighted at music festivals, concerts and competitions around the world. Recently, compositions realized at CCRMA were performed at the International Computer Music Conference in Tokyo; at the Society for Electro-Acoustic Music (SEAMUS) in the U.S., at the Bourges Festival of Electroacoustic Music in France; at ISCM World Music Days; at The Warsaw Autumn Festival in Poland; at the Computers and Music Conference in Mexico City; at the Primera Muestra Internacional de Musica Electroacustica in Puerto Rico; and in concerts in Cuba, Greece, Russia, Argentina, Brazil, Spain, West Germany, Sweden, Switzerland, Italy, Hungary, and Czechoslovakia. Compositions from CCRMA have also won major electroacoustic music prizes over the past few years, including the NEWCOMP contest in Massachusetts, the Irino Prize for Chamber Music in Japan, the Ars Electronica in Germany, and the Noroit Prize in France. Recordings of works composed at CCRMA have been recorded on compact disks by Wergo, Harmonia Mundi and Centaur. CCRMA is publishing with Wergo/Schott "Computer Music Currents", a series of 14 volumes CD's containing computer music by international composers.

Recent compositional works realized at CCRMA include the following:

Ludger Bruemmer – Visiting composer from Essen, Germany.
La cloche sans vallees, and *The Gates of H.* These compositions used sampled sound, sound synthesis, and algorithmic procedures with CLM, CM and RT and a NeXT computer. *The Gates of H.* received first prize in the Ars Electronica 1994.

Joanne Carey – Visiting composer from the U.S.
La Soledad and *Aqui.* Based on the poetry of Pablo Neruda for coloratura and Radio Baton. Composed using DMIX running on a MacIntosh IIfx computer, and an SY77 for sound synthesis.

Chris Chafe – Assoc. Prof. of Music.
Remote Control for MIDI trio. An interactive live performance piece using MIDI instruments and synthesizers; *El Zorro* for brass solo with live electronics. Using NeXT computer, Music Kit, Lightening (Don Buchla's controller), and EMU Proteus; and *El Zorro II* for sea shells and live electronics using the same resources as *El Zorro.*

Kui Dong – DMA Graduate Student.
Flying Apples for tape using DMIX running on a MacIntosh IIfx computer; and *Eclipse I* for tape using CLM on a NeXT computer.

Michael Edwards – DMA Graduate Student.
Redislocations for tape using saxophone samples processed in CLM and Common Music and mixed with RT on a NeXT computer; and *flung me, fóot tród* for alto saxophone and tape using the same resources as above.

RJ Fleck – PhD Graduate Student.
Soundscapes for documentary film and for live performance. Using linear system spectral manipulation of sampled natural sound sources, spoken text and melodic fragments. Resources include real-time signal processing on the Motorola 56000 using software written by Jean Laroche and modified by the composer on a NeXT.

Nicky Hind – DMA Graduate Student.
The Well for 3 female voices and harp. Part of a series of three pieces for female voices plus instrument. Based on ideas from the *I Ching* using algorithmic compositional methods.

David Jaffe – Visiting Composer from the U.S.
Wildlife for Zeta Violin, Mathews/Boie Radio Drum, NeXT Computer and MacIntosh Computer, using the Music Kit, Ensemble (written by the composer) and MAX. A real-time interactive improvisational piece co-composed with Andrew Schloss; *Terra Non Firma* for four cellos and the Radio Baton, using the Mathews Conductor program; and *The Seven Wonders of the Ancient World* a concerto for Radio Drum-driven Disklavier Grand and an ensemble of guitar, mandolin, harp, harpsichord, 2 percussionists, harmonium and bass using MAX.

Peer Landa – Visiting Computer from Norway.
Stroll All Over for computer (direct from disk) and eight strobelights; *Low Motions* for tape and tap-dancer; *Downcast* for tape; and *Irate Sway* for tape and small ensemble. All compositions, except for *Stroll All Over*, were created using signal-processing applications developed by the composer using C on the NeXT computer.

Fernando Lopez-Lezcano – CCRMA Scientific Research Programmer and Composer.
Three Dreams - Paper Castles, Invisible Clouds, Electric Eyes for tape. Composed with sampled sounds and additive synthesis using CLM running on a NeXT computer and a special four channel spacialization unit generator programmed by the composer; and *Espresso Machine* for Radio Drum, NeXT computer and live cello. The composition uses a new improvisational environment built around the Radio Drum. The Radio Drum is used to trigger and control isolated events and event sequences in real-time and communicates through MIDI with the Radio Drum and synthesizers.

Dan Oppenheim – DMA Graduate.
Lamentations for Jerusalem for solo saxophone and DMIX; and *Concerto in D* for violin and DMIX. DMIX was written by the composer specifically for use in composing these pieces and runs on a Mac IIfx computer.

Stephâne Roy – Visiting Composer from Montreal.
Crystal Music for tape. The composition explores the micro-fluctuations of sound, especially those located in the medium and high frequency range and was composed using samples sounds processed using CLM and edited and mixed by traditional studio methods using the Yamaha DMR8 digital mixer. *Crystal Music* won first prize in the Noroit competition in France for 1993.

Atau Tanaka – PhD Graduate Student.
Kagami a real-time interactive piece for performer using a BioMuse controller, MAX, and MIDI synthesizers; *Ets Phon* a virtual guitar piece; and *OverBow* using feedback and FM synthesis control and using the same resources. *Ets Phone* and *OverBow* have been composed at IRCAM where the composer is in residence on a grant from the School of Humanities and Sciences at Stanford University.

Marco Trevisani – Visiting Composer from Italy.
Cosi' e' (from a play by Pirandello). A collaborative piece with the drama department which integrates music, the intermezzo comico and commedia dell'arte with the computer generating music and processing the voices of the actors in real-time. Using SY99, MAX for MacIntosh, CLM and SynthBuilder (see paper on SynthBuilder in following section) for NeXT.

Li Zeng – DMA Graduate Student.
Prelude-song for computer generated tape. Composed using sound sources generated by various techniques. Resources include SY77, MAX, synthesis instruments in CLM, and the Dyaxis-MACMIX package; and *Violin Concerto* for violin and computer tape. Using three sources of sound – CLM, an SY99 controlled by MAX and sampling.

COMPUTER MUSIC SOFTWARE DEVELOPMENT

The CCRMA Music Kit and DSP Tools Distribution

David A. Jaffe
Visiting Composer

Julius O. Smith III
Associate Professor of Music (Research)

The Music Kit is an object-oriented software system for building music, sound, signal processing, and MIDI applications on the NeXT computer. It has been used in such diverse commercial applications as music sequencers, computer games, and document processors. Professors and students have used the Music Kit in a host of areas, including music performance, scientific experiments, computer-aided instruction, and physical modeling. The Music Kit is the first to unify the MIDI and Music V paradigms, thus combining interaction with generality. (Music V, written by Max Mathews and others at Bell Labs three decades ago, was the first widely available "computer music compiler".)

The NeXT Music Kit was first demonstrated at the 1988 NeXT product introduction and was bundled in NeXT software releases 1.0 and 2.0. Beginning with NeXT's 3.0 release, the Music Kit is no longer part of the standard NeXT software release. Instead, it is being distributed by CCRMA via ftp from ccrma-ftp.stanford.edu. Questions regarding the Music Kit can be sent to musickit@ccrma.stanford.edu.

The CCRMA Music Kit and DSP Tools Distribution (or "Music Kit" for short) is a comprehensive package that includes on-line documentation, programming examples, utilities, applications and sample score documents. The package also comes with Bug56, a full featured, window-oriented, symbolic debugger for the DSP5600x signal processing chip family by Ariel Corp.

Source code is available for everything except Bug56 and the low-level sound, DSP, and MIDI drivers. This means researchers and developers may study the source or even customize the Music Kit and DSP Tools to suit their needs. Enhancements can be sent to musickit@ccrma.stanford.edu to have them considered for future CCRMA releases. Commercial NeXT software developers may freely incorporate and adapt the software to accelerate development of NeXTSTEP software products. (Free commercial use of files copyrighted by NeXT Inc. are understandably restricted to NeXTSTEP platforms.)

See the Music Kit Release Notes for further details.

The Music Kit was designed by David A. Jaffe and Julius O. Smith, with input from James A. Moorer and Roger Dannenberg. The Objective-C portion of the Music Kit was written by David A. Jaffe, while the signal processing and synthesis portion was written by Julius O. Smith. The Ensemble application and much of the SynthPatch library were written by Michael McNabb. Douglas Fulton had primary responsibility for the documentation. Others who contributed to the project included Dana Massie, Lee Boynton, Greg Kellogg, Douglas Keislar, Michael Minnick, Perry Cook, John Strawn and Rob Poor.

References

D. Jaffe "An Overview of the NeXT Music Kit," *Proc. 1989 International Computer Music Conference*, Columbus, Ohio, Computer Music Assoc., pp. 135-138.

D. Jaffe and Lee Boynton, "An Overview of the Sound and Music Kits for the NeXT Computer," *Computer Music Journal,* MIT Press, vol. 14, no. 2, pp. 48-55, 1989. Reprinted in book form in *The Well-Tempered Object,* Stephen Pope, ed., 1991, MIT Press.

J. O. Smith, D. Jaffe, and L. Boynton, "Music System Architecture on the NeXT Computer," *Proc. 1989 Audio Engineering Society Conference,* Los Angeles, California, 1989.

M. McNabb, "Ensemble, An Extensible Real-Time Performance Environment," *Proc. 89th Audio Engineering Society Conference,* Los Angeles, California, 1990.

D. Jaffe, "Efficient Dynamic Resource Management on Multiple DSPs, as Implemented in the NeXT Music Kit," *Proc. 1990 International Computer Music Conference,* Glasgow, Scotland, Computer Music Assoc., pp. 188-190.

D. Jaffe, "Musical and Extra-Musical Applications of the NeXT Music Kit," *Proc. 1991 International Computer Music Conference,* Montreal, Canada, Computer Music Assoc., pp. 521-524.

Highlights of the 4.0 Release

- *Conductor Synchronization to MIDI time code & generation of MIDI time code.*
- *Real-time processing of sound from DSP port (with example applications).*
- *Playing sound out the DSP port to external DAT or DAC interfaces, with support for Ariel, Stealth, Singular Solutions and MetaResearch I/O products.*
- *Waveshaping (aka "non-linear distortion") synthesis.*
- *"WaveEdit", the Graphical Waveform editor.*
- *"SynthBuilder", the Graphical SynthPatch editor (separately via ftp).*
- *New substantially-enhanced version of the Ensemble application that supports sound processing instruments, graphical envelope editing, new NoteFilters and more.*
- *New version of the ScorePlayer application that supports playing thorough DSP port.*
- *Support for 32K DSP memory expansion, with automatic sensing.*
- *Support for Ariel QuintProcessor 5-DSP board (with programming examples).*
- *SynthPatch library included as programming example.*
- *Support for quadraphonic sound output via the DSP serial port.*
- *Workspace inspector for scorefiles.*
- *Runs under NEXTSTEP 3.0 and 3.1 on black hardware.*

People who answered the Music Kit survey sent around last year will notice that many of the most-requested items on the survey have been included in the 4.0 release. The obvious missing item is support for NS486 systems. We hope to add this in the coming months. Please send other Music Kit requests to musickit@ccrma.stanford.edu. To subscribe to a Music Kit news group, send email to mkdist-request@ccrma.stanford.edu.

Rapid Prototyping of Digital Signal Processing and Synthesis Systems

Julius O. Smith III, Principal Investigator
Associate Professor of Music (Research)

The nature of computer support for digital signal processing (DSP) research is an ongoing issue. Initially, the Fortran programming language was the standard "high level" representation, and hardware and horizontal microcode served as "low level" representations for mass-produced products. While special purpose hardware (e.g. ASICs) and DSP microcode continue to thrive, still giving the lowest asymptotic cost in mass production, the higher level tools have changed considerably: Fortran is all but obsolete in favor of C, and C is rapidly giving way to its object-oriented extension, C++. For faster research prototyping at the expense of slower execution, interactive programming environments such as MatLab are being used in place of normal software development. These programming environments offer extensive display capabilities and a high-level, interpreted language with easy to use syntactic support of common signal processing operations, both in the time and frequency domains. At a still higher level of abstraction, development tools supporting the direct manipulation of block diagrams are becoming more common. Examples include SimuLink (MatLab), LabView (National Instruments), Ptolemy and Gabriel (Berkeley), Max and TurboSynth (Opcode), SynthKit (Korg R&D), ComDisco, Star, and other CAD tools related to signal processing.

When a rapid prototyping system is designed properly, it is possible to work at all levels in a variety of alternative representations such as block diagrams, object-oriented C, or DSP assembly language.

In typical music synthesis and audio signal processing applications, it is not necessary to sacrifice more than a few percent of theoretical maximum DSP performance, in terms of both speed and code size, in return for the use of a high-level, block-diagram oriented development tool. This is because a small number of primitive modules can implement the vast majority of existing synthesis and processing techniques, and they account for the vast majority of the computational expense. These modules can be fully optimized in advance so that simple, drag-and-drop programming can provide both a real-time simulation and well structured code generation which are very close to the efficiency of a special-purpose, hand-coded, DSP assembly language program. As a result, block-diagram based programming tools are fundamental to good signal processing support in music synthesis and digital audio development systems.

For rapid research prototyping in music and audio applications, there remains an unfulfilled need for a full featured, available, open, and well structured development system supporting MIDI and digital audio synthesis and signal processing. CCRMA is presently supporting the development of **SynthBuilder,** a block-diagram based rapid prototyping tool for these purposes. SynthBuilder was originally based on the program GraSP by Eric Jordan at Princeton University; and GraSP, in turn, was based on SynthEdit by Mike Minnick at NeXT. SynthBuilder has since been developed primarily by Nick Porcaro at CCRMA, with contributions by Julius Smith, David Jaffe, and members of the CCRMA DSP group. SynthBuilder leverages very heavily off of the advanced capabilities of the Music Kit and NeXTSTEP.

References

M. Minnick, "A Graphical Editor for Building Unit Generator Patches," Proc., ICMC-90, Glasgow, pp. 253–255, 1990.

N. Porcaro, J. O. Smith, D. A. Jaffe, and T. Stilson, "The Design of SynthBuilder—A Graphical Synthpatch Development Environment for NeXTSTEP," in preparation, ready by Summer 1994.

The Design of SynthBuilder—A Graphical SynthPatch Development Environment for NeXTSTEP

Nick Porcaro
Visiting Researcher

SynthBuilder is a user-extensible, object-oriented, NeXTSTEP Music Kit application for real-time design of synthesizer patches. Patches are represented by networks of digital signal processing elements called "unit generators" and MIDI event elements called "note filters" and "note generators". SynthBuilder is based on Eric Jordan's GraSP application and was developed using the Music Kit and object oriented design.

A distinction is made between event-level processing, which is handled by note generators and note filters, and signal processing, which is handled by unit generators. The link between the two types of processing is made with special type of note filter called the PatchParameter which provides mapping of events to unit generator control methods. This processing is managed using the object networking mechanisms provided by the Music Kit.

The semantics of unit generator parameter control are encapsulated by high-level classes which can have sophisticated graphical interfaces. Such class design makes it easy to support the notion that synthesis and processing are two sides of the same coin; sound input is just another unit generator, as is sound output.

The real-time support of the Music Kit means there is no "compute/then listen" cycle to slow down the development of patches. This is possible because SynthBuilder actually uses two processors at the same time; the DSP and the host processor. While the DSP is being used to synthesize sound the host processor is free to support the graphical user interface, which can modify the sound while it is being processed on the DSP.

Dynamic loading is extensively used. DSP code is linked and loaded while sound is produced, custom inspectors are loaded dynamically only when needed, and note-filters/note-generators are linked dynamically. This enables convenient user extensibility, and the patch-edit response time, application launch time, and development turn-around time are extremely short.

SynthBuilder has features similar to the version of MAX that runs on the IRCAM Signal Processing Workstation, but offers compatability with the MusicKit, and only requires the built-in DSP56001 chip.

SynthBuilder generates Music Kit SynthPatches for Music Kit applications, which is analogous to how InterfaceBuilder creates interfaces for NeXTSTEP applications; hence, SynthBuilder can be viewed as an interface-builder/instrument-builder for the DSP. In addition, SynthBuilder generates working inspector code that is more immediately useful than the stub code generated by InterfaceBuilder.

SynthBuilder can potentially generate code for other systems, and thus is a cross-compiler of sorts. Finally, the high-level block diagram representation of a patch network is potentially more portable and easier to understand than assembly language, C or Objective-C.

Common Lisp Music and Common Music Notation

William Schottstaedt
Research Associate

Common Lisp Music (CLM) is a music synthesis and signal processing package written primarily in Common Lisp. It is aimed at composers who are not primarily interested in computer-related performance issues or computer assisted improvisation – there is no support for "real-time" interactions.

The instrument design language is a subset of Common Lisp (its numerical functions, and nearly all its control functions), extended with a large number of "unit generators": Oscil, Env, Table-Lookup, and so on. The "run-time" portion of an instrument can be compiled into C, 56000, or lisp code. Since CLM instruments are lisp functions, a CLM note list is just a lisp expression that happens to call those functions. The notes need not even be in any order. The actual computation is done one note at a time, and the results are overlayed, building up the sound note by note. If an Ariel Quint Processor board is available, CLM can take advantage of it, computing many notes in parallel.

Common Music Notation (CMN) is a music notation package written in Common Lisp, using CLOS and the Sonata font. It provides for all the usual needs of music notation in a fully customizable, programmable environment.

CLM and CMN have recently been extended, ported to new machines, and more tightly integrated with Rick Taube's Common Music program. In CLM, new features include a fast mixing package, a "make" facility for sound files, support for AIFF and other sound file formats, ports to the NS/Intel platforms and new Macs, and several new instruments. In CMN, we have continued to add various specialized notations. Common Lisp Music and Common Music Notation are available free, via anonymous ftp [pub/Lisp/clm.tar.Z and pub/Lisp/cmn.tar.Z at ccrma-ftp.stanford.edu].

References

Taube, H. "COMMON MUSIC - A Music Composition language in Common List and CLOS," *STAN-M-63,* Stanford University, January 1990.

A Dynamic Spatial Sound Movement Toolkit

Fernando Lopez-Lezcano
Scientific Research Programmer

This brief overview describes a dynamic sound movement toolkit implemented within the context of the CLM software synthesis and signal processing package.

The goal of the project is to provide a set of tools that can be used by a composer working within the CLM environment to represent, control and render multiple moving sound sources in, for now, bidimensional space. The user interface consists of a unit generator for controlling the spatial movement and the environment in which it occurs (dlocsig), and a set of CLOS classes for the description in simple terms of the path each sound object follows (path). The composer describes the movement by specifying a small set of points in space and the software routines generate interpolated Bezier curves that create a smooth trajectory between the specified points.

The low level hardware support for four channel reproduction is an external high quality four channel digital to analog converter box (the QuadBox) connected to the DSP port of a NeXT workstation. The corresponding software support consists of a C/DSP56000 assembler program that can play standard four channel soundfiles through the QuadBox.

Part of the work in this project was done as a team effort with Atau Tanaka at CCRMA while Fernando was working at the SFC Campus, Keio University, Japan.

Common Music / Stella: A Music Composition Language in Common Lisp and CLOS

Heinrich Taube
Visiting Composer

Common Music (CM) is a compositional language implemented in Common Lisp and CLOS and takes advantage of CLOS's multiple inheritance and method combination to support a wide range of compositional entities and associated behaviors. Because of the underlying representation, Common Music encourages user extensions to both its classification scheme (either by "mixing in" preexisting classes or defining new root classes from scratch) and to the predefined instances of objects that the system uses. Fundamental to the language is a classification taxonomy that defines compositional concepts such as scores, parts, scales, notes, rhythms, dynamics and compositional tools such as item streams, functions, algorithms and generators. A rich and complex compositional environment is built from these basic concepts through subclassification and mixins. Common Music system currently runs on NeXT, Macintosh, SGI, SUN and 486/Windows computers and outputs music in the following formats: realtime MIDI and MIDI file, Common Lisp Music sound file and score file, Common Music Notation Postscript and input file, MusicKit score file, CSound scorefile, and RT score file.

Stella is a system that uses persistent objects to support algorithmic and editorial styles of music composition in Common Music. By algorithmic it is meant the specification of musical material is made indirectly through a program that creates it. By editorial it is meant the specification of musical material is made directly by the composer. Using Stella's representation system, the composer may select or mix either style as appropriate to a given compositional situation. Stella comes with an editor that supports a simple, non-Lisp command interface. This editor treats a composition as a structured data base and supports a powerful query language for the algorithmic examination and modification of musical material.

[Rick Taube is currently on the research staff of the Institut fuer Musik und Akustik at the Zentrum fuer Kunst und Medientechnolgie in Karlsruhe, Germany, where he is continuing work on this project. He can be reached via email at: hkt@zkm.de.]

References

Schottstaedt, B.. "Pla: A Composer's Idea of a Language," *Computer Music Journal,* 7(1):11-21, 1983.

DMIX

Daniel V. Oppenheim
DMA Graduate
currently at Computer Music Center
IBM T.J. Watson Research Center
Yorktown Heights, NY 10598

1. Composition and performance are but a part of the creative musical process. This is especially true when composing with the NEW musical materials made available by computers.

2. Composition of new music must be an experimental process in which the composer attempts to implement his musical ideas. Quickly experiencing new materials and ideas within the overall musical context is the only way to really evaluate them. A system for composition should enable to easily try out new musical ideas as they formulate. Furthermore, the flow of creative ideas should not be affected by the need to deal with technical aspects of the system.

3. Expressivity, nuances, and detail may be as important as the musical idea itself. Whereas in instrumental music the composer can rely on a performer and performance practice to add the 'spice of life' to his music, in computer music the composer must have tools that will enable him to control and express any amount of musical detail.

4. Graphic manipulation, computer programming, improvisation, real-time interaction, and off line composition, are all valid ways for working with music. A system should make it easy to constantly move between diverse modes of working and tools as the musical context constantly evolves.

QUILL is a full fledged algorithmic composition language, sharing many concepts with PLA and Common Music. Quill offers several input formats, including the full power and support of the Smalltalk and DMIX object systems. Quill supports an MLF (Multi Level File) file format enabling to split a document into a hierarchy of pages that can be worked on individually. It is interactive in the sense that it can talk directly to graphic objects and views, interact with real time input objects (Echos), and offers a full fledged interactive debugger.

Graphical Editing views support piano roll notation, parameter views, function views, filter views, hierarchy views of large sections of music, and more. A Graphic view can be Slapped into Quill, where it will be transformed into a program that will generate the same music. Graphic Views also support real-time editing, that allows the user to hook any parameter to an external MIDI control (slider, button, etc.), modify the parameter as it is playing, and get both audio and visual feedback. Selections allow the user to manipulate non contiguous selections of music events, both graphically and via programming.

Echo is a class of interactive real time objects, much like MAX patches. Echos integrate into the entire DMIX environment and can be easily extended by the user via trivial Smalltalk programming.

SHADOW is an interactive performance system. Shadow can track a live performer, match the tempo and dynamics of the accompaniment, and be used to trigger any DMIX or Smalltalk event in response to any performer gesture or input. The combination of Shadow and Echo allow both traditional and non-traditional score tracking techniques.

Besides the above DMIX supports numerous tools for manipulating music, such as Functions, Filters, Patterns, and more, tools for saving and restoring any DMIX objects from disk (including MIDI files), tools for navigating through the environment, on-line interactive, context sensitive, and user editable help, and much more.

PHYSICAL MODELING AND DIGITAL SIGNAL PROCESSING

Digital Waveguide Modeling of Acoustic Systems

Julius O. Smith III
Associate Professor of Music (Research)

Digital Waveguide Filters (DWF) have proven useful for building computational models of acoustic systems which are both physically meaningful and efficient for applications such as digital synthesis. The physical interpretation opens the way to capturing valued aspects of real instruments which have been difficult to obtain by more abstract synthesis techniques. Waveguide filters were initially derived to construct digital reverberators out of energy-preserving building blocks, but any linear acoustic system can be approximated using waveguide networks. For example, the bore of a wind instrument can be modeled very inexpensively as a digital waveguide. Similarly, a violin string can be modeled as a digital waveguide with a nonlinear coupling to the bow. When the computational form is physically meaningful, it is often obvious how to introduce nonlinearities correctly, thus leading to realistic behaviors far beyond the reach of purely analytical methods.

A basic feature of DWF building blocks is the exact physical interpretation of the contained digital signals as samples of traveling pressure waves, velocity waves, or the like. A byproduct of this formulation is the availability of signal power defined instantaneously with respect to both space and time. This instantaneous handle on signal power yields a simple picture of the effects of round-off error on the growth or decay of the signal energy within the DWF system. Another nice property of waveguide filters is that they can be reduced in special cases to standard lattice/ladder digital filters which have been extensively developed in recent years. One immediate benefit of this connection is a body of techniques for realizing any digital filter transfer function as a DWF. Waveguide filters are also related to "wave digital filters" (WDF) which have been developed primarily by Fettweis. Waveguide filters can be viewed as providing a discrete-time "building material" incorporating aspects of lattice and ladder digital filters, wave digital filters, one-dimensional waveguide acoustics, and classical network theory. Using a "mesh" of one-dimensional waveguides, modeling can be carried out in two and higher dimensions.

In this context, a *waveguide* can be defined as any medium in which wave motion can be characterized by the one-dimensional wave equation. In the lossless case, all solutions can be expressed in terms of left-going and right-going traveling waves in the medium. The traveling waves propagate unchanged as long as the wave impedance of the medium is constant. The wave impedance is the square root of the "massiness" times the "stiffness" of the medium; that is, it is the geometric mean of the two sources of resistance to motion: the inertial resistance of the medium due to its mass, and the spring-force on the displaced medium due to its elasticity.

Digital waveguide filters are obtained (conceptually) by sampling the unidirectional traveling waves which occur in a system of ideal, lossless waveguides. Sampling is across time and space. Thus, variables in a DWF structure are equal *exactly* (at the sampling times and positions, to within numerical precision) to variables propagating in the corresponding physical system. In other applications, the propagation in the waveguide is extended to include frequency dependent losses and dispersion. In still more advanced applications, nonlinear effects are introduced as a function of instantaneous signal level.

When the wave impedance changes, signal scattering occurs, i.e., a traveling wave impinging on an impedance discontinuity will partially reflect and partially transmit at the junction in an energy-preserving way. Real-world examples of waveguides include the bore of a clarinet, a violin string, horns, organ pipes, the vocal tract in speech, microwave antennas, electric transmission lines, and optical fibers.

References

J. O. Smith, "Physical Modeling Using Digital Waveguides," *Computer Music Journal,* special issue: Physical Modeling of Musical Instruments, Part I, vol. 16, no. 4, pp. 74–91, Winter, 1992.

Music 421 (EE 266) Course Description, J. O. Smith, Stanford University.

Waveguide Reverberators and Real-Time Implementations

R. J. Fleck
PhD student, Music

Efficient digital simulation of reverberation characteristics can be obtained through the use of waveguide filters as an alternative to other methods. Work previously done at CCRMA with conventional reverberation techniques on earlier computer systems and still in use today (i.e. the 'nrev' instrument found in CLM) has not been improved upon for some time.

Some of the potential advantages regarding waveguide filters in the context of reverberation include: 1) a correlation to physical room-acoustic behaviour which is flexible and allows for configurations of arbitrary complexity 2) good and predictable control over the artifacts of quantization and round-off error and 3) an exact control over total system energy.

Some of the goals of this current research area are to discover: 1) the potentials of recent technological advances in the area of reverberation and room acoustics (ie. waveguides, updated analysis/synthesis approaches, etc.) 2) the possilibities for improved implementation at CCRMA and 3) problems inherent in the implementation of a graphically-interactive interface for real-time simulations.

References

M. R. Schroeder,"Natural Sounding Artificial Reverberation," *JAES,* 10:3 p.219, July 1962.

J. A. Moorer,"About this Reverberation Business," *Computer Music Journal,* 3:2 p.13., 1979.

J. O. Smith,"A New Approach to Digital Reverberation using Closed Waveguide Networks," in "Music Applications of Digital Waveguides," Department of Music Report No. STAN-M-39, 1987.

J. O. Smith,"Music Applications of Digital Waveguides," Department of Music Report No. STAN-M-39, 1987.

J. M. Jot, and Chaigne, A "Digital Delay Networks for Designing Artificial Reverberators," *Proc. 90th AES Conv,* preprint 3030, 1991.

J. M. Jot, "An Analysis/Synthesis Approach to Real-Time Artificial Reverberation," *IEEE,* 0-7803-0532-9 1992.

Feedback Delay Networks

Davide Rocchesso
Visiting Scholar
PhD Student - Universita' di Padova

Recursive comb filters are widely used in signal processing, particularly in audio applications such as digital reverberation and sound synthesis. In the recent past, some authors [Stautner-Puckette '82, Jot '92] have considered a generalization of the comb filter known as the multiple feedback delay network (MFDN). The main purpose of this research is to investigate some efficient implementations of MFDN's and to find some interesting applications.

The MFDN is built using N delay lines, connected in a feedback loop through a set of scattering coefficients. These coefficients may be organized into a "feedback matrix". If such a matrix is unitary, system poles have magnitude one and the MFDN has only constant-amplitude eigenmodes. For the structure to be practically useful, an attenuation coefficient must be applied at the output of each delay line to adjust the length of the impulse response.

D. Rocchesso has proposed restricting the feedback matrix to a circulant structure. The resulting Circulant Network (CN) allows an easy computation of its eigenvalues from Fast Fourier Transform of a row (or a column). Using equal-length delay lines, we can exactly compute the frequencies where resonances occur, even for very large matrices. Moreover, by using nearly-equal-length delay lines, the decay time can be controlled with a single coefficient. This structure is also proper for VLSI implementation because it can be efficiently made parallel.

It has been shown how to use CN's for many purposes in sound processing and synthesis: for simulation of radiating structures such as instrument bodies, for simulation of feedback resonators, and even for live electronics performances. These possibilities extend the range of applicability of MFDN's beyond reverberation.

CN's with short delay lines may be used to produce resonances irregularly distributed over frequency. A possible application could be the simulation of resonances in the body of a violin. In this application the exact position and height of resonances are not important. By changing delay lengths, it is possible to move poles in frequency, while by changing the network coefficients we can re-shape the frequency response. The loop gain determines the maximum peak to valley distance. Such a structure using short delay lines has been used in live-electronic-sound processing, where a dynamic filtering can be achieved by changing the MFDN parameters in real time.

CN's are also very effective as resonators in Karplus-Strong-like algorithms, especially for simulating membranes or bars.

Connections between MFDN's and Digital Waveguide Networks (WGN) (Smith '85) have been revealed. Julius O. Smith and D. Rocchesso have shown that the FDN is isomorphic to a (normalized) waveguide network consisting of one (parallel) scattering junction and N branches, each connecting to the one scattering junction at one end, and reflectively terminated at the other. This correspondence gives rise to new generalizations in both cases. Theoretical developments in this field are in progress.

References

A. Paladin and D. Rocchesso, "Towards a Generalized Model of One-Dimensional Musical Instruments," *Proc. 1992 Int. Workshop on Models and Representations of Musical Signals*, Capri Italy, Universita' di Napoli Federico II, 1992.

A. Paladin and D. Rocchesso, "A Dispersive Resonator in Real-Time on MARS Workstation," *Proc. 1992 Int. Conf. Computer Music*, San Jose California, Computer Music Assoc., 1992.

D. Rocchesso and F. Turra "A Generalized Excitation for Real-Time Sound Synthesis by Physical Models," *Proc. 1993 Stockholm Music Acoustics Conference,* Stockholm Sweden, Royal Institute of Technology (KTH), 1993.

D. Rocchesso "Multiple Feedback Delay Networks for Sound Processing," *Proc. 1993 X Colloquio di Informatica Musicale,* Milano Italy, Universita' degli Studi di Milano, 1993.

D. Rocchesso and J. O. Smith "Circulant Networks for Sound Synthesis and Processing," Submitted to Int. Conf. Computer Music, Aarhus Denmark, Computer Music Assoc., 1994.

J. O. Smith and D. Rocchesso "Connections between Feedback Delay Networks and Waveguide Networks for Digital Reverberation," Submitted to Int. Conf. Computer Music, Aarhus Denmark, Computer Music Assoc., 1994.

The 2–D Digital Waveguide Mesh

Scott Van Duyne
PhD student, Music

Julius O. Smith III
Associate Professor of Music (Research)

The traveling wave solution to the wave equation for an ideal string or acoustical tube has been modeled efficiently with bi-directional delay-line waveguides. Two arbitrary traveling waves propagate independently in their respective left and right directions, while the actual pressure at any point may be obtained by summing the theoretical pressures in the left– and right–going waves.

Excellent results have been obtained modeling strings and acoustic tubes using one-dimensional waveguide resonant filter structures. However, there is a large class of musical instruments and reverberant structures which cannot be reduced to a one-dimensional traveling wave model: drums, plates, gongs, cymbals, wood blocks, sound boards, boxes, rooms—in general, percussion instruments and reverberant solids and spaces.

In the two dimensional case of wave propagation in an ideal membrane, the traveling wave solution involves the integral sum of an infinite number of arbitrary plane waves traveling in all directions. Therefore we cannot just allocate a delay line for every traveling plane wave. Finite element and difference equation methods are known which can help with the numerical solution to this problem; however, these methods have had two drawbacks: (1) their heavy computational time is orders of magnitude beyond reach of real time, and (2) traditional problem formulations fit only awkwardly into the physical model arena of linear systems, filters, and network interactions.

Our solution is a formulation of the N-dimensional wave equation in terms of a network of bi-directional delay elements and multi-port scattering junctions. The essential structure of the two-dimensional case is a layer of parallel vertical waveguides superimposed on a layer of parallel horizontal waveguides intersecting each other at 4-port scattering junctions between each bi-directional delay unit. The 4-port junctions may be implemented with no multiplies in the equal impedance case. Plane waves, circular waves, and elliptical waves all propagate as desired in the waveguide mesh. Band limited accuracy can be enforced. The three-dimensional extension of the waveguide mesh is obtained by layering two-dimensional meshes and making all the 4-port junctions into 6-ports, or though a tetrahedral, four-port, no-multiply structure.

The two-dimensional waveguide mesh is mathematically equivalent to the standard second-order-accurate finite partial difference formulation of the wave equation. It, therefore, exhibits the desirable stability and convergence properties of that formulation. However, the numerical solution methods of initial value problems involving second order hyperbolic partial difference equations usually require a multi-step time scheme which retains values for at least two previous time frames. The waveguide mesh reduces this structure to a one-step time scheme with two passes: (1) In the computation pass, the scattering junction computations are performed in any order (a feature well–suited to parallel computation architectures); then (2) in a delay pass, their outputs are moved to the inputs of adjacent junctions.

Current work on the waveguide mesh is in (1) exploring alternative spatial sampling methods, (2) developing efficient hardware implementation structures, (3) introducing loss and dispersion into the mesh in a physically correct, yet efficient, manner, and (4) finding the right parameters to model specific musical instruments and spaces.

The Felt Mallet

Scott Van Duyne
PhD student, Music

Julius O. Smith III
Associate Professor of Music (Research)

Recent work has led to digital waveguide string models and physical models of membranes using a 2–D digital waveguide mesh. We are currently working on ways to excite these models in physically correct ways. One obvious need is a good model of the felt mallet for drums and gongs, and of the piano hammer for strings.

The attack transient of a struck string or membrane can be approximated by the injection of an appropriate excitation signal into the resonant system. However, this excitation method is not sufficient to cope with the complexities of certain real musical situations. When a mallet strikes an ideal membrane or string, it sinks down into it, feeling a pure resistive impedance. In the membrane case, the depression induces a circular traveling wave outward. If the membrane were infinite, the waves would never return, and the mallet would come to rest, losing all its energy into the membrane. If the membrane is bounded, however, reflected waves return to the strike point to throw the mallet away from the membrane. The first reflected wave to reach the mallet may not be sufficiently powerful to throw the mallet all the way clear, or may only slow down its motion, and later reflected waves may finally provide the energy to finish the job. This complex mallet-membrane interaction can have very different and difficult to predict acoustical effects, particularly when a second or third strike occurs while the membrane is still in motion.

In our model, we view the felt mallet as a nonlinear mass/spring system, the spring representing the felt portion. Since the felt is very compliant when the mallet is just barely touching the membrane, yet very stiff when fully compressed, we must use a nonlinear spring in the model, whose stiffness constant varies with its compression. Our essential requirements are that the model be digitally efficient, that it be easily interconnected to waveguide structures, and that it be able to compute arbitrarily accurate strike transients from measured data from real hammers, strings, mallets, and drums.

The Haptically Controlled Waveguide Piano

Julius O. Smith
Associate Professor of Music (Research)

Scott Van Duyne
PhD Student, Music

Brent Gillespie
PhD Student, Mechanical Engineering

et. al.

Making a good piano synthesis algorithm traditionally has not been easy. The best results thus far have been in the area of direct sampling of piano tones. This approach is memory intensive, as there is a lot of variety in the piano timbre ranging from low notes to high, and from soft to loud. In addition, sampling techniques don't have a good answer to the problem of multiple strikes of the same string while it is still sounding, nor to the coupling of strings which are undamped while other strings are sounding. We believe that the solution will be found through waveguide modeling and DSP synthesis techniques. An even more intrinsic problem of synthesizers is that they don't feel anything like a piano when you play them. Currently at CCRMA, we have the good fortune of having a variety of people working separately on solutions to different parts of the piano problem, although their individual work may have broader applications.

The piano problem may be broken down into five basic parts: (1) the string, (2) the soundboard, (3) the piano hammer and damper, (4) the key mechanism itself, and (5) the implementation hardware and software.

The primary difficulties of modeling the string are found in that the piano string harmonics are not exactly harmonic, and in that there is significant coupling between horizontal, vertical and longitudinal modes on the string. In addition, there may be important nonlinear effects. Work being done by Julius Smith on string coupling and fitting loop filters to measured data will solve some of the problems. Other work by Scott Van Duyne and Julius Smith will lead to simplifications in modeling the stretching harmonics of the piano string. It may be that work relating to passive nonlinearities by John Pierce and Scott Van Duyne will be helpful for the finest quality tone generation.

The soundboard can now be modeled in a fully physical way using the 2-D Digital Waveguide Mesh, a recent development by Scott Van Duyne and Julius Smith which extends waveguide modeling techniques into two or more dimensions. Julius Smith is working on applying results from his work on bowed strings to piano synthesis; extremely efficient new algorithms are possible using this approach.

The excitation of waveguide string models has till now been left primarily to loading the string with an initial condition and letting it go, or to driving the waveguide loop with an excitation signal tailored to the desired spectral response of the attack transient. While almost any attack transient may be achieved through driving the model with an excitation signal, the variety of interactions that a piano hammer may have with a string is immense when one considers the possibilities ranging over the very different high and low strings, and over the wide range of strike forces. Further, it would be virtually impossible to catalog the possible attack transients due to a hammer hitting a string which is already in motion due to a previous strike. The hammer/string interaction is very complex. Fortunately, recent work initiated by Scott Van Duyne and continued with Julius Smith on modeling felt mallets as wave digital nonlinear mass/spring systems will allow all these complex interactions to fall directly out of the model.

Brent Gillespie's work on the touchback keyboard project will provide a realistic controlling mechanism for the piano synthesis algorithm. The touchback keyboard is a haptic control mechanism, driven by a computer controlled motor. It looks like a piano key, and feels like a piano key. That is, it senses the force applied by a person to the key, and computes, in real time, the correct key motion response based on the equations of motion of the internal key mechanism. It is easy for the touchback keyboard to provide the felt hammer element of the tone synthesis algorithm with a hammer strike velocity. This velocity will be used to drive the synthesis algorithm. In return, the piano hammer element can provide the touchback keyboard with a return hammer velocity at the right time, and the person playing the key will feel the appropriate haptic response.

The hardware and software to implement this complete piano model is available now. The touchback keyboard is controlled by a PC with an add-on card dedicated to real-time computations of the equations of motion. The NextStep operating system running on a Next or PC platform will provide a suitable environment for the synthesis algorithm. Specifically, the SynthBuilder Application being developed by Nick Porcaro and Julius Smith provides a cut-and-paste prototyping environment for real-time DSP-based audio synthesis, and the Music Kit, being maintained and improved by David Jaffe, provides higher-level access to the DSP 56000 card.

There is additional research interest in vibrotactile feedback in the piano keys as suggested in the current work of Chris Chafe. While this effect may be less important in the modern piano, it is certainly more important in early keyboard instruments, and critical in the clavichord, where the hammer may remain in contact with the vibrating string after striking it. Further, we shall want to make the piano sound as if it were somewhere in a particular room or concert hall. Work by John Chowning in localization, by Jan Chomyszyn in loudness perception, by Steven Trautmann in speaker arrays, and by R. J. Fleck in efficient reverberation models can round off the final auditory experience.

Synthesis of the Singing Voice Using a Physically Parameterized Model of the Human Vocal Tract

Perry R. Cook
Research Associate

Two voice synthesis systems have been constructed using a physically parameterized model of the human vocal tract. One system is a real-time Digital Signal Processor (DSP) interface, which allows graphical interactive experimentation with the various control parameters. The other system is a text-driven software synthesis program. By making available both real-time control and software synthesis, both rapid experimentation and repeatable results are possible. The vocal tract filter is modeled by a waveguide filter network. Glottal pulses are stored and retrieved from multiple wavetables. To this periodic glottal source is added a filtered pulsed noise component, simulating the turbulence which is generated as air flows through the oscillating vocal folds. To simulate the turbulences of fricatives and other consonants, a filtered noise source can be made arbitrarily resonant at two frequencies, and can be placed at any point within the vocal tract. In the real-time DSP program, called SPASM, all parameters are graphically displayed and can be manipulated by using a computer mouse. Various two-dimensional maps relating vowels and vocal tract shapes are provided, and a user can smoothly vary the control parameters by moving the mouse within a map region. Additional controls include arbitrary mapping of MIDI (Musical Instrument Digital Interface) controls onto the voice instrument parameters. The software synthesis system takes as input a text file which specifies the events which are to be synthesized. An event specification includes a transition time, shape and glottal files as written out by the SPASM system, noise and glottal volumes, glottal frequency (either in Hz. or as a musical note name), and vibrato amount. Other control strategies available include text-to-speech/singing and a graphical common music notation program. Support for languages, musical modes, and vocal ornamentations is provided in Latin, and modern Greek.

Adding Pulsed Noise to Wind Instrument Physical Models

Chris Chafe
Associate Professor of Music (Research)

Pulsed noise has been detected in the residual of steady flute tones after elimination of purely periodic components. LMS adaptive linear periodic prediction was used to track the waveform through its slight period-to-period fluctuations. The predicted signal was removed from the original leaving a breathy sounding residual to examine. Noise pulses in musical oscillators result from period synchronous gating of the driving means. Bowed string instruments exhibit noise pulses arising from alternating stick-slip motion, where noise is introduced only when the string is slipping. Distinct pulses are also exhibited by the saxophone in which the reed modulates air friction. Flute noise is more continuous than in string or reed tones. Short time fourier transformation of the residual signal reveals that pulses are present, but spectrally weighted toward higher frequencies. A physical model of the flute incorporating a corresponding noise synthesis method is being developed. Results of the simulation are compared for quality of pitch synchronous spectral modulation and effect on frequency jitter.

The method uses a vortex-like noise generator mechanism coupled to the nonlinear excitation mechanism. These components simulate the flute's frictional noise generation and switching air jet, respectively. The vortex is generated by a separate short-cycle nonlinear oscillator. It's output is used to modulate the nonlinearity of the main instrument (for example, a cubic polynomial in Perry Cook's SlideFlute). The vortex's signal input is a flow variable which is controlled by the signal circulating in the main instrument loop.

The resulting oscillation contains noise injected by the vortex and exhibits the desired pitch synchronous spectral changes. The possible classes of instruments for which this might apply include air jet, glottis, single, double and lip reed.

Physical Modeling of Brasses

David Berners
PhD student, Electrical Engineering

One of the difficulties in building waveguide models of brasses and winds is that we do not know how to find the round-trip filtering in a flaring horn without actually making an acoustic measurement. Ideally, we would like to be able to compute the loop filter directly from the physical dimensions of the horn. While significant work has been done along these lines (Causse et al. [1], Plitnik and Strong [2], Benade [3]), a complete and accurate theory is not yet available.

To provide computationally tractable models, the flaring horn is modeled assuming that Webster's horn equation is satisfied, i.e., that a one-parameter solution to the wave equation exists within the boundaries of the horn. Any shape, such as planar or spherical, can be assumed for the wavefront within the horn.

In an ongoing research project at CCRMA, Webster's horn equation is solved as follows: First, the wave equation is converted to the form of the celebrated Schrodinger wave equation through a coordinate transformation outlined by Benade in [3]. Once in Schrodinger's form, the wave equation becomes equivalent to the one-dimensional scattering problem in particle physics, for which efficient and numerically stable solution methods exist (Kalotas and Lee [4]). In the new (transformed) coordinate system, the horn boundary function is replaced by the "horn potential function", which, in addition to providing the frequency dependent reflection, transmission, and impedance functions for the waveguide, can be used to gain an intuitive understanding of how these characteristics are related to bell flare. The quantities obtained from the solution to Webster's equation are all that is necessary for the design of lumped filters to be used in a digital waveguide model. Advantages over conventional modeling techniques include the ability to specify an arbitrary wavefront shape and possible numerical advantages.

References

Causse, R. et al "Input impedance of brass musical instruments - Comparison between experiment and numerical models," *J. Acoust. Soc. Am.*, Vol. 75, No. 1 p. 241

Plitnik, G. R. and Strong, W. J. "Numerical method for calculating input impedances of the oboe," *J. Acoust. Soc. Am.*, Vol. 65, No. 3, March 1979, p. 816

Benade, A. H. "The Physics of Brasses," *Scientific American*, 24-35 July 1973 p. 23

Kalotas, T. M. and Lee, A. R. "A new approach to one-dimensional scattering," *American Journal of Physics*, 59 (1) Jan 1991 P. 48

Synthesis and Research on Reed Driven Woodwind Instruments with Particular Emphasis on the Saxophone

Gary P. Scavone
PhD student, Music

The modeling of musical instruments using digital waveguide techniques has proven to be both an accurate and efficient technique for synthesis. Because such models are based on physical descriptions, they further provide a useful platform for acoustical explorations and research. Realistic clarinet models have been developed (Cook), as well as an interactive development environment (Hirschman) for the study of reed dynamics. To date, the clarinet has drawn most of the attention, mainly because of its relative simplicity. Models of the saxophone have been scarce, though the recently released Yamaha VL1 demonstrates that the saxophone too can be effectively modeled using waveguide techniques.

A simple saxophone model has been designed and implemented in the NeXTStep programming environment. Though still in early stages of development, an interactive platform is planned which will facilitate the study of questions regarding human/saxophone coupling and mouthpiece influences. Such explorations should further result in accurate synthesis models for saxophones.

A Passive Nonlinear Filter for Physical Models

John R. Pierce
Visiting Emeritus Professor of Music (Research)

Scott A. Van Duyne
PhD student, Music

Nonlinearities, small or large, favorably affect the sounds of many musical instruments. In gongs and cymbals, a gradual welling-up of energy into the high frequencies has been observed. Nonlinearities cause the transfer of energy from lower modes to higher modes after the instrument has been struck. These nonlinearities do not generate new energy, only transfer it. While memoryless square–law and look–up table nonlinearities may be incorporated in computer generation of sounds, these means often cause system energy loss or gain, and are difficult to control when a range of large and small effects are desired.

Our approach to the injection of nonlinearity into resonant systems was to identify a simple passive nonlinear electrical circuit, and then to apply physical modeling techniques to bring it into the digital signal processing domain. The result was an efficient digital nonlinear mode coupler which can be attached to any waveguide termination, or inserted into any resonant digital system where traveling waves are being computed. The mode coupler can be tuned to set the rate of energy spreading as well as the region of the spectrum to be affected. Excellent results have been obtained creating gong and cymbal crash sounds by connecting these passive nonlinear filters to 2–D Digital Waveguide Mesh boundary terminations.

Much work remains to be done mapping measured nonlinear effects in real musical instruments on to parameters of our filters.

Physical Modelling of Music Instruments Using Neural Networks

Wen-Yu, Su
Visiting Scholar

Physical modelling of music instruments has been studied at CCRMA for years. A Digital Waveguide network may create a successful model for music instrument synthesis by simulating the wave propagation in the instrument. However, finding the correct input signals and internal filter parameters to a specific Waveguide instrument model still requires further investigation.

A new method using Neural Networks is proposed to solve the above problems. When the squashing functions in Neural Networks are linear, Waveguide networks can be regarded as a subset of Neural Networks. Through the training of a Neural Network, it is possible to design a synthesizer for a specific instrument as long as the input/output relationship is given and the shape of the network imitates the instrument physically. The other advantage of using Neural Networks is that the nonlinear properties of an instrument can also be modelled when it is necessary.

In order to use Neural Networks as physical modelling tools for music instrument synthesis, the following problems have to be addressed: (1) Mechanical analysis of instruments must be carried out in advance to collect the relevant data; (2) new protocols for Neural Networks have to be designed according to the new requirements; and (3) dedicated training algorithms have to be developed which take into account stability problems and physical constraints.

FFT-Based Signal Processing and Spectral Modeling Synthesis

Julius O. Smith III, Principal Investigator
Associate Professor of Music (Research)

The Fast Fourier Transform (FFT) revolutionized signal processing practice in the 1960s. Today, it continues to spread as a practical basis for digital systems implementation. Only very recently has it become cost-effective to use the short-time FFT in real-time digital audio systems, thanks to the availability of sufficiently powerful, low-cost, single-chip solutions.

In the digital audio field, FFT-based techniques are ripe to appear in digital mixing consoles, post-production editing facilities, and top-quality digital audio gear. Many music and digital audio "effects" can be conveniently implemented in a unified way using a short-time Fourier analysis, modification, and resynthesis facility.

In the music synthesis field, obtaining better control of sampling synthesis will require more general sound transformations. To proceed toward this goal, transformations must be understood in terms of what we hear. The best way we know to understand a sonic transformation is to study its effect on the short-time spectrum, where the spectrum-analysis parameters are tuned to match the characteristics of hearing as closely as possible. Thus, it appears inevitable that sampling synthesis will migrate toward spectral modeling. Recent developments in constant-Q filterbanks, such as in the wavelet literature, have created new alternatives for consideration. Advanced time-frequency representations, such as the Wigner Distribution, are yielding new insights into time-varying audio spectra.

In contrast with physical modeling synthesis which models the *source* of a sound, spectral modeling techniques model sound at the *receiver,* the human ear. Spectral modeling is more immediately general than physical modeling since it is capable of constructing an arbitrary stimulus along the basilar membrane of the ear, while new physical models must be developed for each new class of musical instrument. While complex coarticulation effects are more naturally provided by physical models, the short-time Fourier transform can be applied to any sound demonstrating any desired effect to determine what must happen in a spectral sequence to produce that effect.

FFT-based techniques play an important role in (1) the practical implementation of general signal processing systems (fast convolution), (2) advanced effects such as "cross synthesis," time compression/expansion, duration-invariant frequency shifting, and other "phase vocoder" type techniques, and (3) novel synthesis systems based on the direct creation and transformation of spectral events and envelopes.

References

Serra, X., and J. O. Smith, "Spectral Modeling Synthesis: A Sound Analysis/Synthesis System Based on a Deterministic plus Stochastic Decomposition," *Computer Music Journal,* vol. 14, no. 4, pp. 12–24, Winter, 1990.

Music 420 (EE 265) Course Description, J. O. Smith, Stanford University.

Instantaneous and Frequency-Warped Signal Processing Techniques for Pitch Tracking and Source Separation and Fast Linear-Phase FIR Filter Theory and Design

Avery Wang
PhD student, Electrical Engineering

In the past year I have been engaged in basic signal-processing research. My main achievements were continued work on the problem of instantaneous frequency tracking, insights into frequency-warped signal processing, harmonic signal separation, and a fast linear-phase FIR filtering algorithm and design methodology.

Last year I presented some preliminary results of my basic Frequency-Locked Loop tracking algorithm at the CCRMA Associates meeting. Since then I have made progress in that area. I have developed a modification using the Kay window to achieve minimum variance instantaneous frequency estimates.

By using an array of FLL partial trackers it is possible to track time-varying harmonic series by forming an averaged estimate of the fundamental frequency from the trackers and constraining the tracker center frequencies to be integer multiples of the fundamental. Forming a constrained net of trackers has the advantage of minimizing the estimate variance of the instantaneous fundamental frequency. This technique has many potential uses in speech parameterization, speech compression, pitch shifting, and voice separation. I am currently writing a NeXTStep application to demonstrate these ideas.

Analysis of frequency-varying signals is generally problematic using conventional techniques such as Fourier transforms. Using dynamic frequency demodulation it is possible to separate signal bandwidth into components due solely to frequency modulation and to envelope amplitude modulation. By factoring signals into AM and FM components it is possible to convert a high-variance (Fourier) signal model to a low-variance model consisting of an instantaneous frequency track and an envelope function, thus minimizing the signal's spread in the time-frequency plane and facilitating signal understanding and mixture separation. I call this signal processing point of view "frequency-warped" signal processing due to the time-varying frequency demodulation applied to the input signal.

Finally, I developed a theory of fast FIR filtering which may reduce the computational complexity of many FIR filtering tasks by a factor of ten or more. Virtually any desired filter magnitude response may be specified and a corresponding fast linear-phase filter may be designed for it. The main idea is to truncate an IIR impulse response and create a filter with the reversed impulse response. The two filters cascaded thus yield a linear-phase filter. I have specified design procedures, provided a statistical analysis of error accumulation and performance limits, and written code in Matlab to aid filter design. Additionally, I have specified a technique for performing fast, minimum-variance frequency estimation based on the Kay window. These algorithms have many uses in audio technology where fast linear-phase filtering is desired, for example in digital mixing boards and digital loudspeaker crossover systems. A patent disclosure has been filed with the Stanford Office of Technology Licensing.

Time-Frequency Analysis of Audio

Scott Levine
PhD student, Electrical Engineering

Current methods in signal analysis are dependent upon the Short-Time Fourier Transform (STFT) for mapping discrete-time signals to a discrete-time discrete-frequency domain. But, the constant dimensions of this time-frequency plane tiling is not well suited to current models of the human ear. The human ear can be crudely modeled as a filterbank whose bandpass filters have constant-Q. That is, the ratio of their center frequency to their bandwidths are constant over all filters. This is not the case for the STFT; using the Fourier transform gives equal time and frequency bandwidth to all filters, regardless of their center frequency.

By the Heisenberg uncertainty principle, one cannot have arbitrarily small bandwidth in both time and frequency. Their product has been shown to have a lower bound; so improved resolution in one domain decreases resolution in another domain. To optimize this trade-off, I propose a time-frequency analysis such that all frequencies will have resolution up to a semitone in pitch. Since pitch is on a logarithmic scale, low frequencies will have higher resolution than high frequencies. Accordingly, high frequencies will have better time resolution than low frequencies. It is this non-uniform tiling of the time-frequency plane that sets it apart from traditional STFT methods. This new tiling will also be better matched to the transfer function of the human ear, and will thus be able to convey more useful information.

This time-frequency tiling is constructed from a combination of the Discrete Wavelet Transform (DWT) and the Fast Fourier Transform (FFT). The DWT analysis is implemented in a critically sampled, octave-band iterated filterbank. Then, the outputs of these filterbank channels are analyzed separately with the FFT. The combination of these two fast algorithms (they can both be implemented in $O(N \log N)$ complexity) create a bounded-Q frequency analysis. Frequency is first tiled into separate octaves, which are logarithmic in frequency, with the use of the DWT. Then with the use of the FFT, the frequency resolution is constant within an octave. This approximation to the constant-Q model of the ear can have many new applications from improved equalization, to polyphonic music transcription, to source separation.

CONTROLLERS FOR COMPUTERS AND MUSICAL INSTRUMENTS

Real-time Controllers for Physical Models

Chris Chafe
Associate Professor of Music (Research)

Perry R. Cook
Research Associate

The computational reductions brought about by new work in algorithms such as the Waveguide filter formulations, along with improvements in DSP chips and other signal processing hardware, have made possible the real-time synthesis of music by physical modelling. Such instruments require new modes and levels of control. Work in increasing the bandwidth of synthesizer control by exploiting all available degrees of freedom has yielded a number of experimental hybrid controllers (Cook, Chafe). Controllers based on the paradigms of wind and stringed instruments have improved the control of models based on these families, and research is being conducted to create a more general controller which is not constrained to a particular family of instruments.

Mapping physical gestures to a DSP synthesis controller is being studied by experimentation. Early studies in simulation (Chafe, 1985) suggested that linear mappings are not the way to go. The current development system allows trial-by-feel investigation of alternative scalings.

The area of tactile feedback (Chafe) is being investigated, as this is an important area of control for the traditional instrument player. Initial trials have begun using actuators feeding audio to the touch point. A general preference has been shown with the technique. The next stage will be to quantify what enhancement, if any, results from feeling the instrument's vibrations. Also, such considerations as tactile frequency bandwidth and vibrations characteristic of contact points will be studied.

New pieces are being written using real-time controllers and the DSP-based physical models. "El Zorro" is a recent composition by Chris Chafe that employs a Lightning Controller (by Buchla and Associates). The soloist is steering note-generation algorithms in terms of tempo, tessitura and "riff-type." Gesture and position is tracked with the Lightning's infra-red controllers. Some direct control is exercised over DSP effects via MIDI. A composition project in the works uses the Celletto (an electronic cello) to interact with the DSP synthesis at the control level. The cellist will evoke synthesis related to natural cello behavior directly from the instrument. For example, bow speed might translate into breath pressure control of a wind synthesis.

Ongoing Work in Brass Instrument Synthesizer Controllers

Perry R. Cook
Research Associate

Dexter Morrill
Visiting Composer, Colgate University

Brass instrument players have been at a disadvantage when using their instruments as computer music controllers, because they have been limited to commercial pitch extractors which do not measure and use the unique spectral and control features of the brass instrument family. In this project, brass instruments were fitted with several sensors and were used in conjunction with specially designed pitch detection algorithms.

Systems were constructed using various valved brass instruments. Pressure sensors located in the mouthpiece, on the bell, and in mutes are used for pitch detection and pickup of the direct horn sound. Switches and linear potentiometers were mounted near the valves for finger control, and traditional foot pedals are also available for control. Optical sensors were mounted on the valves, providing information about valve position.

The valve and acoustic information are used to form pitch estimates which are both faster and more accurate than those yielded by commercial pitch extractors. The other switches and controls provide control over MIDI synthesizer parameters such as sustain, patch change, and controller changes, as well as controlling signal processor parameters.

Performer-Oriented Brass Instrument Synthesis and Control

Timothy Stilson
PhD Student, EE

I am exploring ways of bringing electronic music performance to brass players. Currently, there is very little a brass player can do if he/she wants to take advantage of the world of electronic music, other than taking up another instrument. Current brass controller systems use standard pitch detection schemes that can be rather slow or error-prone. Perry Cook investigated adding parameters such as key position to the algorithm to give accurate initial guesses and narrower search spaces for the controller. Chris Chafe experimented with tactile feedback in other musical areas as a possible aid to performance. I am investigating ways to extend those ideas.

The main reason brass controllers haven't been as easy to construct (or play) as more well-known reed controllers is that the player's lips play a much more integral part in the production of the sound. In a reed instrument, the vibrations occur at the reed, whose vibration parameters are influenced by the lips, bite, and breath of the performer, all of which are relatively simple to measure. In a brass instrument, the reed is replaced by the lips themselves. The player has quite a bit of control over the vibration parameters of the lips such that measurement is nearly impossible. Thus, instead of measuring the parameters that control the reed vibration, one can merely guess the lips' parameters based on the the output of the horn (or the sound inside the horn). This results in either a time delay or an error in the output of the controller. Controllers that try to make a quick guess based on measurable values have the common problem of register ambiguity: based on measurable parameters (valve positions, breath pressure, etc.), it is nearly impossible to tell which of several notes will be played (the actual note is determined mostly by the player's lip tension, which is one of the hard-to-measure quantities).

I am investigating various ways of lessening the difficulties associated with the brass controller. I am attacking the problem from several angles: 1) I am constructing an algorithm that could produce quicker output from an instrumented brass instrument; 2) I am exploring possibilities for actually measuring certain lip parameters so that note ambiguity can be vastly decreased; and 3) I am developing a vibration-feedback controller that should allow the player to play much more intuitively without requiring that he play a real instrument, whose physical sound output may be unwanted in certain performance situations.

If feasible, the vibration-feedback controller, which makes use of waveguide synthesis to simulate the correct feedback vibrations will be a very versatile instrument. Since a complete brass instrument model is necessary to simulate the feedback vibrations, we get for free the output of the brass model. Thus, the controller should be able to act as a brass instrument itself. In fact, if one views this not as a controller, but as a *brass instrument simulator*, one sees some interesting results:

1. The instrument can be made to simulate the brass instrument so closely that the lip feedback interface could actually become nearly indistinguishable from a real instrument from the player's perspective. Note, the player would play the controller just like a real instrument: blowing into it and vibrating his lips, the feedback vibrations would interact with the lip vibrations just like the resonances of a real horn, so the phenomenon of mode-locking would occur (certain frequencies being easy to play and others impossible to play). This phenomenon is the most basic feedback a player uses when playing a horn, and it is very difficult to produce artificially in any usable way. This is the problem I am tackling in making the simulator.

2. This is the most compelling result of the above property: if well crafted, the simulator could be used as a brass instrument in performance. As long as the lip interface acts just like a real instrument would (from the performer's perspective), and as long as the output sounds realistic, the simulator should be essentially no different in practice from a physical brass instrument, and it could be played as such.

3. Since the simulator is under real-time program control, simulating different instruments is as simple as loading a new program (and changing mouthpieces, since these can't be simulated and are vital to the performance characteristics of the instrument). Thus one could conceivably simulate *any* brass instrument, given a well-designed lip interface. The possible uses of the simulator are many: it can replace multiple instruments in certain situations, like schools and pit orchestras; be a tunable instrument (Bb, C, C# trumpet, anyone?); or simulate extinct brass instruments or impossible ones.

4. The most difficult part of a brass-instrument waveguide-synthesis model is modeling the lips and their interaction with the pressure waves in the bore. The simulator can ignore that part of the model since it actually happens physically in the lip interface.

5. One drawback: such a simulator gains little advantage in trying to extract performance parameters for controlling other synthesizers. Since the model is physically based, all values and parameters in the simulator are fundamentally no different than those one would be able to get from attaching sensors to a real instrument. Luckily, many of those parameters are in a much more accessible form (i.e. they would be hard to get to in a real instrument, like in the middle of the bore of the valve section of the trumpet, for example, or maybe the sensor would interfere with the sound or operation of the instrument). Thus the same algorithms necessary for extracting performances from real instruments would be necessary on the simulator, just a little easier to implement.

Finally, as a side project, I am exploring ways for a whistler to perform with more sounds than just the whistle sounds, yet maintaining the flexibility of pitch (and other parameters) available in whistling. This involves a mixture of audio processing, DSP, and pitch extraction to produce more complex and interesting sounds based on a relatively pure input signal.

References

Perry Cook, Dexter Morrill, and Julius O. Smith, "A MIDI Control and Performance System for Brass Instruments," *Proc. 1993 Int. Conf. Computer Music*, Tokyo, Japan, Computer Music Assoc., 1993.

Chris Chafe, "Tactile Audio Feedback," *Proc. 1993 Int. Conf. Computer Music*, Tokyo, Japan, Computer Music Assoc., 1993.

Haptic User Interfaces for the Blind

Sile O'Modhrain
Visiting Scholar

Brent Gillespie
PhD student, Mechanical Engineering

Advances in graphic output technology have opened the window to the development of advanced graphical user interfaces making computers increasingly inaccessible to the blind. To date, developers seeking to overcome this situation have relied on two methods of outputing information - sound and braille. Neither of these have been able to provide an adequate substitute for graphics. For certain applications, such as synthesizer controllers and digital music editing, speech output would conflict with the audio output of the system. Therefore we feel it is necessary to explore other ways of presenting information in a tactile form. Because haptic displays can, like graphics, create virtual objects, they present a more natural analogue than text (as in speech or braille). For example, a motorized mouse can define a button to be felt as well as seen—imagine that a particular area of the mouse pad has a different texture. This force reflecting system would also usefully supplement graphical user interfaces for sighted users.

With a simple powered joystick built from spare parts we have produced such effects as a virtual button and a variety of texture cues. Those who have used this haptic prototype agree that we have begun to tap a very promising resource.

We aim to develop an application for the PC which will be capable of making the graphic output of commercial sound editing programs accessible by touch. To this end, we are writing routines to read and recognize icons from video memory. We can already read arbitrary buttons from the screen and display them through sound (midi) and touch (powered joystick).

The Touchback Keyboard

Brent Gillespie
PhD student, Mechanical Engineering

A virtual piano action (a mathematical model), a simulation algorithm, and a set of 88 single degree-of-freedom haptic display devices constitute a promising means to make a synthesizer keyboard feel like a real grand piano. Other keyboard instruments could be simulated with the same device at the touch of a button. We propose, then, to simulate the feel of the grand piano by numerically integrating the equations of motion of the multibody piano action in real-time in a human-in-the-loop simulation scheme. As the integration proceeds, the finger-key interaction force is computed and generated by a haptic interface. In addition, the motion (needed for use in the integration) is sampled from the haptic interface hardware.

We believe that haptic (tactile/kinesthetic) feedback from an instrument, in addition to the aural feedback, plays a crucial role in the process by which a keyboardist controls the tone and timing. Further experiments in human-computer interaction can be made with arbitrarily determined dynamical behaviors of the keyboard.

The behavior of the simulated piano action should appear passive to the user even though it is realized with active devices. We seek design guidlines which will ensure that such dyanamical behaviors characteristic of active objects (chatter and stickyness) will not arise. Applications of both linear and non-linear analytical tools are being applied to this end.

The Stanford Radio Drum

Max V. Mathews
Professor of Music (Research)

There are two new designs for the radio drum. They are simpler than the existing Boie drum and thus are believed to be more suitable for manufacture. All these drums sense the positions of any number of drum sticks in three dimensional space, x, y, and z.

The first design uses only four electrodes as the receiving antenna rather than the approximately 100 electrodes in the Boie design. Also, the design uses completely separate circuits to sense x and y, thus simplifying the data processing and possibly improving the accuracy of the measurements. The antenna can be fabricated as a large, single, two-sided printed circuit whose dimensions are the size of the drum head. The second design uses five electrodes for the receiving antenna, each of which is essentially a rectangle. The electrodes are simple enough so they can be made from sheet metal or even aluminum foil. Consequently, the cost of making the large printed circuit can be saved.

The second design also has the advantage that with a suitable computer program to interpret the data, it can be used with arbitrarily shaped and positioned electrodes. Thus, it is conceivable to attach electrodes to an existing musical instrument and sense the gestures which the performer uses to play the instrument.

The New Conductor Program

Max V. Mathews
Professor of Music (Research)

A new version of the conductor program has been written. It performs all the functions of the original program and has additional features. The main advantages of the new program are simplicity and greater speed. Also the new program can be easily modified or extended.

The main difference in the structure of the new program concerns the way in which multiple voices are handled. In the original program, the notes in each voice were stored in time order in separate lists. During the performance, the lists were merged by a real-time sorting operation done concurrently with playing the notes. In the new program, the voices are merged ahead of time into one time-ordered list by the score compiling program. This single list is then read and the notes played by the conductor program itself. Eliminating the real-time sorting operation substantially speeds the performance program. It also simplifies the program so that more elaborate tempo control algorithms can be used.

The score in the original program contains only pitches and durations (in beats) for each note. The new score can have additional information attached to the notes, for example accents or parameters affecting timbre.

In the old program, all voices in a given score had to be written for the entire duration of the score even though they might be tacit for long periods of time. In the new program, voices need only be written for the actual blocks when they are playing notes.

PadMaster: An Improvisation Environment for the Radio Drum

Fernando Lopez-Lezcano
Scientific Research Programmer

PadMaster is a real time improvisation environment that is controlled through the Radio Drum. It was written in the NextStep environment using Objective-C, NeXT's Interface Builder for the graphical user interface and the MusicKit for the basic MIDI and music performance class library. PadMaster receives through MIDI position and force information from the Radio Drum when either of the batons hits the surface of the drum and can also poll continuously for their instantaneuos spatial position.

The program enables the composer to define virtual pads in the drum surface and assign a behavior to each one (a 5 by 6 grid of rectangular pads in the current implementation). The pads can be grouped in sets and the performer can step forward or backwards through the sets using the batons, in effect completely (or subtly) changing the behavior of the drum and the interaction of the performer with the composition. Each pad can trigger the next note in a sequence of notes; trigger, pause and resume a single sequence or trigger multiple overlapping instances of a sequence (each sequence can be one note, or a list of notes or any other MIDI messages). Each pad can be controlled by a global tempo or can have its own tempo, or even each sequence within a pad can have a completely independent and variable tempo. In addition each pad can request the main controller program to send continuous position information and can translate that information into tempo control or continuous MIDI messages such as controllers or pitch bend. The graphical interface provides feedback to the performer as to the state of each individual pad (idle, playing, paused and so on). All the information can be saved and loaded from document files, edited by using graphical inspectors or loaded and saved as text in MusicKit score files. Future work will enable the program to use additional MIDI controllers in addition to the Radio Drum.

The Computer-Extended Ensemble

David A. Jaffe
Visiting Composer

Until recently, there have been two basic models of how electronics interact with a performer in a performance situation. One model adds a tape of synthesized sound to an instrumental ensemble. We call this the "tape music" model. The other model, "keyboard electronic music", consists of pianists performing on keyboard synthesizers. In the case of the tape music model the performer is forced to slave to the electronics; with keyboard music it is the electronics that slave to the performer. We are beginning to realize that these two models are actually end points of a continuum, with the region between them largely unexplored.

The central question for composers is not whether human behavior can be duplicated, but what new musical effect can be achieved with computer interaction that cannot be achieved by prior existing means? A likely place to begin exploring this question is in an area of music in which interaction between performers is central, improvisation.

Introducing a computer as an extension of the improvising performer increases the scope of the kind of spontaneous musical decision-making that give improvisational music its distinctive quality. A computer can magnify, transform, invert, contradict, elaborate, comment, imitate or distort the performer's gestures. It gives the performer added power to control sound at any scale, from the finest level of audio detail to the largest level of formal organization.

But the full power of the computer in an improvisational context does not show itself until we add a second performer to the ensemble. Now each performer can affect the playing of the other. One performer can act as a conductor while the other acts as soloist. Both performers can be performing the same electronic instrument voice at the same time. And these roles can switch at a note-by-note rate. Thus, the walls that normally separate performers in a conventional instrumental ensemble become instead permeable membranes. Figuratively speaking, the clarinetist can finger the violin and blow the clarinet while the violinist bows the violin and fingers the clarinet. We have coined the term "computer-extended ensemble" for this situation.

The challenge becomes finding roles for the performers that allow them just the right kind of control. They need to feel that they are affecting the music in a significant and clear fashion. Otherwise, they will feel superfluous and irrelevant, as if the music has gotten out of control. The computer program may be simple or complex, as long as it fires the imagination of the performers.

We have been experimenting in this realm with percussionist/composer Andrew Schloss in a duo called "Wildlife." The duo features Schloss and the author performing on two modern instruments, the Mathews/Boie Radio Drum and the Zeta electronic/MIDI violin, with this ensemble extended by two computers, a NeXT and a Macintosh running the NeXT Music Kit and Max. The music is a structured improvisation in which all material is generated in response to the performers' actions; there are no pre-recorded sequences or tapes.

Current work include "The Seven Wonders of the World", unlike "Wildlife", casts the computer and Radio Drum in the context of a conventional ensemble (or, at least, a conventionally notated ensemble.) This piece is scored for Radio Drum-controlled Disklavier, harpsichord, harp, two percussionists, mandolin, guitar, harmonium and bass. It was composed at the Banff Centre for the Arts, where I was a Visiting Artist in 1992-93.

Other projects include the following:

1. "Terra Non Firma", a work for conducted electronic orchestra and four cellos, using the Matthews Conductor program. This work was commissioned by the University of Victoria in honor of Max Matthews.

2. "American Miniatures", a recently completed work for tape alone, uses Common Music, the Music Kit and the phase vocoder to process recorded sounds of strings, voices and drums.

Real-Time Control Using Biological Signals

Bill Putnam
PhD student, Electrical Engineering

The focus of this research effort is to use naturally occuring bioelectric signals in advanced human computer interfaces. The specific signals used to date include the Electroencephlogram (EEG), Electroocularggram (EOG) and Electromyogram (EMG). These three signals arise from the normal electrical activity associated with neural activity in the body. The EEG signal is the electrical activity arising from brain activity, EOG signals come from the eye positioning mechanism, and EMG signals are the electrical signals produced by the activity of the musclature.

Biomuse is a real-time system used to collect, analyze and utilize biological signals for control purposes. Biomuse was developed at CCRMA by Ben Knapp and Hugh Lusted . Currently, this system accomodates the collection 8 channels of biological data. Biomuse operation is controlled by host software running on a PC. Communication between the host and the Biomuse takes place through a standard serial interface. In addition, the Biomuse includes a MIDI interface to directly connect to musical instruments.

My work to date has focused on the use of the EMG signal. To this end, pattern recognition techniques have been used to detect and classify dynamic gestures on the part of the user. A typical example would be that of determining if the user was opening or closing their hand. Once classified, these gestures have been used to effect volitional control in the context of a computing environment and musical performance.

There are two basic application areas which are being pursued. The first is that of musical performance. As mentioned previously, pattern recognition techniques can be used to classify physical gestures made by the user. Once classified, these gestures can be used for real-time control of a performance. Work in this area is being focused upon the development of a real-time mixing system which would respond to both EMG and EOG signals. Knapp has succesfully implemented an eye tracking device using EOG signals. The eye tracker could be used to select individual instruments by simply looking at them. Once selected a parameter such as volume could then be controlled by an appropriate gesture, such as moving the arm up or down.

The second application area is that of computer control. Current efforts have focused upon cursor control and control of standard graphical objects such as sliders, scroll bars etc. The main focus and end use of this effort is to facilitate human computer interaction for persons with disabilities.

*This work has been supported by Biocontrol Systems Inc.

Related publications

Patmore, David and William Putnam. "Assistive Cursor Control For A PC Window Environment: Electromyogram and Electroencephlogram Based Control," *Proceedings of the Virtual Reality and Persons With Disabilities Conference,* June 1994.

Putnam, William and R. Benjamin Knapp. "Real-Time Computer Control Using Pattern Recognition of the Electromyogram," *Proceedings of the Engineering in Medicine and Biology Conference,* pp.1236-1237, 1993.

Putnam, William and R. Benjamin Knapp. "The Use of the Electromyogram in a Man-Machine Interface," *Proceedings of the Virtual Reality and Persons With Disabilities Conference,* June 1993.

Biocontrol Interfaces as Musical Instruments

Atau Tanaka
PhD student, Music

The BioMuse is a neural interface/biocontroller developed by BioControl Systems, Palo Alto, CA. It monitors electrical activity in the body (as EEG and EMG) and translates it into MIDI [Knapp and Lusted 1990]. I have been composing music for and performing with the BioMuse. Although such human interface technology can be applied in ways to give the capability of making music to nonmusicians, my interest has been to look at the system as a new musical instrument - one that requires "practice" to avoid playing "wrong notes" in performance, and an instrument for which we can imagine developing an idiomatic performance technique [Tanaka 1993].

MIDI output from the BioMuse is passed to software running in the Max environment [Puckette and Zicarelli 1990]. These Max patches configure the BioMuse, and also map incoming MIDI control data (representing EMG trajectories) to musical gestures. In this way, a physical gesture of the muscles effects melody, rhythm, timbral changes, and combinations.

There is a certain frustration in directly connecting the BioMuse output to MIDI devices in this way. The source biodata is a rich, continuous signal that is constantly changing. MIDI, on the other hand, is an event based music control specification. To better suit the nature of the biosignal, I have created Max patches to allow direct control of sound synthesis by sending MIDI System Exclusive to the synthesizer.

This idea has been expanded upon at IRCAM by the use of the IRCAM Signal Processing Workstation (ISPW). Here, synthesis and signal processing algorithms are created directly in an expanded version of the Max program. In this set up, output from the BioMuse can directly affect changes in sound production without the limitations of MIDI synthesizers.

Another area of exploration has been the control of visual images from the BioMuse. Animation and video files are created in the Macintosh as PICS files and QuickTime movies. Max patches are then created to navigate through these dynamic graphics images under control of the BioMuse. This allows one bodily gesture to pilot at once changes in the music and visuals. I feel that this holds interesting potential for multimedia performance.

References

Knapp, Benjamin R. and Hugh S. Lusted. "A Bioelectric Controller for Computer Music Applications," *Computer Music Journal,* 14(1):p.42, 1990.

Puckette, Miller and David Zicarelli. *MAX - An Interactive Graphic Programming Environment,* Opcode Systems, Palo Alto, CA, 1990.

Tanaka, Atau. "Musical Technical Issues in Using Interactive Instrument Technology," *Proceedings of the International Computer Music Conference,* Tokyo, 1993.

PSYCHOACOUSTICS AND COGNITIVE PSYCHOLOGY

Applying Pychoacoustic Phenomenon to the Coordination of Large Speaker Arrays

Steven Trautmann
PhD student, Music

Since the earliest loudspeakers, there has been a trend to create systems that "sound better" than their predecessors. While this has included all aspects of recording and reproduction of sound, a significant aspect has been using speakers in a coordinated fashion to give the illusion of localized sound sources. The work of Chowning, Bosi, and Lopez-Lezcano at CCRMA has resulted in an effective means of creating illusory localized sources on four or more channel systems. However, further improvements can be made by controlling the exact timing and content of the signals sent to speakers at known locations in a known environment since the pressure fluctuations in space are then a determined quantity.

By using many speakers, increasingly better simulations are possible. In general, N speakers can exactly reproduce a signal generated by a localized sound source at N selected points in space if the speakers are coordinated properly. This coordination is done by solving a system of linear equations to design the proper filters. By using a least squares or some other appropriate method, the pressure fluctuations at a large number of points can be approximated with a smaller number of speakers. Choosing these points in a spatial sampling pattern allows calculation of how many speakers are needed for a given a situation and error limit over the entire space. Room acoustics can cause problems, so the system would work better in a non-reverberant environment.

Trade-offs between magnitude error and phase error are also possible, allowing further refinements utilizing psychoacoustics. It is hoped that by using masking phenomenon, the perception of phase and magnitude as well as head and ear acoustics, good results will require fewer speakers and therefore potentially less overall processing.

Distance of Sound in Reverberant Fields

Jan Chomyszyn
PhD student, Music

The creation of convincing auditory perspective is an important element of computer music; it makes the sound lively and expressive. Many factors contribute to the impression of space and the location of sound sources, including appropriate reverberation, and balance of loudness and timbres of the sounds used in the composition. Some of the parameters which provide cues to distance of the sound sources are correlated in a natural reverberant environment. A typical example is direct-to-reverberant sound energy ratio and intensity, which change reciprocally along the physical distance between the sound source and the listener.

However, the percepts arising from the physical cues do not always follow the same relationship. This is easy to show in the visual world, in the case of size constancy. In visual perspective, to preserve the impression of size constancy of an object, the physical size of the object has in fact to be diminished in proportion to the provided perspective. Is this also the case in auditory perspective? Existing evidence seems to confirm this thesis.

Since the beginning of this century researchers are aware that changes in loudness and changes in distance may sometimes form equivalent concepts for the listeners (Gamble 1909). As a part of his "physical correlate theory" Warren noticed that loudness judgements of his stimuli (speech) depended on the degree of reverberation (Warren 1973). Recently Chowning (Chowning 1990) observed that loudness constancy takes place in room environment in an analogous way to the size constancy in vision. An experiment being carried out investigates this postulate with regard to computer music.

Dry percussive sounds have been produced by using the physical model of a hammer (Van Duyne 1994), simulating varying effort of the player. The sounds have been next added reverberation in the rate corresponding to changing distance in a room. In the test, subjects match dry prototypes to each of the reverberated sounds. They are also being suggested the auditory perspective of the room before each trial. Care has been taken to eliminate possible influence of the spectral bandwidth on the loudness match. The test will reveal if distant sounds played with greater effort are perceived as louder, as if the loudness were estimated at the sound source.

References

Chowning, J.M. *Music form Machines: Perceptual Fusion & Auditory Perspective - for Ligeti.* Report No. STAN-M-64, CCRMA Department of Music, Stanford Universiy, 7-14, 1990.

Gamble, E.A. McC. Minor studies from the psychological laboratory of Wellesley College, I. Intensity as a criterion in estimating the distance of sounds. *Psychological Review* , 16, 416-426, 1990.

Van Duyne, S. personal communication, 1994.

Warren, R.M. "Anomalous loudness function of speech," *Journal of the Acoustical Society of America,* 54, 390-396, 1973.

Pitch and Repetition Rate Perception

John R. Pierce
Emeritus Visiting Professor of Music(Research)

There are many demonstrably strange examples of musical pitch. Risset produced a tone whose pitch falls a little when the tape speed is doubled. If one alternates a sine wave and a tone with many harmonics in going down a scale from the top to the bottom of the piano keyboard, the sine wave simply becomes inaudible near the bottom of the keyboard; for low notes pitch simply cannot be determined from the fundamental. If one goes down a scale to the bottom of the piano keyboard alternating a tone with successive harmonic partials with a tone with odd harmonic partials only, toward the bottom end the tone with odd harmonic partials has a higher pitch than the tone with successive harmonic partials. The simplest explanation of this and other behavior is that pitch involves recognition of patterns of vibration along the basilar membrane. For high musical pitches the fundamental or first partial is of predominant importance in judging pitch. In the mid-range the fundamental is important if it is there, but we still recognize the pattern of harmonics and the pitch if the fundamental is absent. For very low pitches, the fundamental is of little importance, and we recognize pitch from the pattern of harmonics. This is in accord with how pianos are tuned.

Experiments have been carried out studying a related phenomenon. Sequences of tonebursts (sinusoids turned smoothly on and off) have a clear pitch if they are some tens of cycles long. If such tonebursts are only 2 to 4 cycles long they are heard as dull or bright clicks, depending on toneburst frequency. Sequences of such tonebursts with different relative phases can have different ratios between fundamental frequency and number of tonebursts per second. At low rates (below 300 tonebursts per second for 4800 Hz tonebursts) sequences with the same number of tonebursts per second sound the same. At considerably higher rates, sequences with the same fundamental frequency sound the same. Results obtained so far seem best explained by a basilar membrane place mechanism of pitch perception and some other competitive, but consistant pitch in pitch, mechanism for judging rate. The most relevant work on toneburst sequences was completed and published in "Periodicity and pitch perception," *Jour. Acoust. Soc. Am.*, vol 90, pp. 1889-1893, October 1991.

Musical Perception and Memory

Daniel J. Levitin
Visiting Research Fellow, CCRMA
University of Oregon, Dept. of Psychology

I am studying the nature and accuracy of human auditory memory, in particular, memory for musical events. When people hear a musical piece, what aspects of that piece are preserved in memory, and with how much accuracy?

The first phase of this research examined the memory of non-musicians for musical pitch. 25% of the people tested sang popular songs from memory in the correct pitch; 67% of the people came within two semitones of the correct pitch. This suggests that memory for pitch is stable, and that a large number of people possess some component of what we traditionally think of as absolute pitch.

In the second phase of this research, I am examining the extent to which people have internalized our western, A-440 based scale. What we call scale tones are, of course, just arbitrary markers along a logarithmic frequency continuum. When people sing from memory, will they tend to sing tones that cluster around these arbitrary markers or will the tones they produce form a somewhat random distribution along the continuum? Preliminary results indicate that there is indeed a clustering, at least for some people, around the markers of our A-440 scale. This suggests that, through exposure, people have internalized this scale as a stable standard. Such a finding is consistent with neurological evidence that throughout every stage of the auditory system there exist cells that are selectively tuned to particular frequencies. It seems plausible that this tuning is affected by repeated exposure to sounds in the surrounding environment.

In the third phase of this research, I am studying the stability of memory for tempo. Preliminary analysis suggests that when people sing songs from memory, they tend to sing those songs in—or very near—the proper tempo for those songs.

Taken together, these studies provide evidence that (1) memory for perceptual events is more stable and accurate than previously recognized; (2) human memory systems code and preserve many of the features of musical stimuli, and perhaps perceptual stimuli in general; (3) the well-known ability of "Absolute Pitch," and the less known ability of "Absolute Tempo" may exist to some degree in even non-musicians with no special training.

Statistical Models for Psychoacoustic Research

Daniel J. Levitin
Visiting Research Fellow, CCRMA
University of Oregon, Dept. of Psychology

Physical quantities such as musical pitch and color are commonly represented as linear frequencies, with low frequencies at one end of a line and high frequencies at the other. Yet the perception of musical pitch is better characterized by a circular structure: frequencies group into octaves, so that tones an octave apart are perceived as more similar than tones less than an octave apart. The same is true of color: the colors at the extremes of the spectrum, red and violet, are perceived as more similar to each other than to those in the middle of the spectrum.

The statistical methods commonly used for linear quantities are inappropriate for circular structures and may lead to serious statistical errors. A branch of statistics, known as circular statistics, was introduced by Lord Rayleigh at the turn of the century to accomodate just these cases. However, the methods of circular statistics have remained relatively obscure and have not yet become commonplace in psychological and psychophysical laboratories.

I am working to expand the knowledge and use of circular statistical methods by:

1. preparing demonstrations of the errors that can easily be made by improperly using linear statistics on circular data;
2. compiling tables of critical values for common circular statistical goodness-of-fit tests;
3. running Monte Carlo simulations to determine the relative power and effectiveness of various circular test statistics;
4. creating computer programs to perform circular statistical tests. None exist to date.

Psychological Representation of English Vowel Sounds

Roger N. Shepard
Professor of Psychology

Perry R. Cook
Research Associate

Daniel J. Levitin
Visiting Research Fellow, CCRMA
University of Oregon, Dept. of Psychology

Human speech is composed of elementary perceptual units called phonemes. These are the smallest units that we recognize as separately producible speech sounds, the vowels and consonants of our alphabet. Great effort has been focussed over the last three decades on designing sophisticated devices for the recognition and production of speech. A better understanding of how humans perceive speech might lead to advances in computer speech systems, in addition to expanding our knowledge of the human mind.

This research focusses on how humans perceive English vowel sounds, and we hope to "map" the psychological structure of vowel sounds in a way that reveals this underlying structure. The "map" will be a multi-dimensional space in which vowels that are perceived as similar will appear close together, and those that are perceived as dissimilar will appear far apart. Previous solutions have been propsoed based on confusion data (Peterson and Barney; Shepard), articulatory models of speech perception (Liberman) and acoustic or spectral analysis of the speech signal.

In a new approach, we are employing an apparent motion paradigm to order the vowel sounds in similarity space.

Apparent motion is a perceptual illusion that occurs when stationary objects appear to move, due to limitations in the processing speed of the brain. For example, suppose we have two lights that are a particular distance from one another, and that flicker on and off at a particular rate. An illusion of motion is created, wherein one light appears to be moving to the position of the other. This phenomenon is the basis of motion pictures (which are comprised of a sequence of still frames) and of movie theater and Las Vegas nightclub marquis (wherein lightbulbs seem to form a moving pattern).

Apparent motion unfolds according to a predictable schedule of stages:

1. At slow alternation rates between the light sources, no motion is perceived, and the lights appear simply to alternate.
2. At more rapid alternation rates, apparent motion is perceived.
3. At still more rapid alternation rates, apparent motion breaks down, and a different perceptual event occurs, known as visual stream segregation. During visual stream segregation, people perceive the two lights as flashing on an off rapidly in two separate "streams"; no motion is perceived between them, and they do not seem to alternate.

It has been known for nearly a century that the alternation rates required for apparent motion to occur - and the alternation rates at which it breaks down into stream segregation - are linearly related to the distance between the objects. It is this relation between time and distance (Korte's Third Law of Apparent Motion) that provides the rationale for the present study.

By presenting pairs of vowel sounds in rapid alternation, we will discover the time at which steam segregation occurs. This can be used as a measure of distance between the vowel sounds.

Preliminary results suggest that the technique is valid, and that a representation of the psychological structure of speech sounds—at least according to this method—is attainable. What remains to be seen is whether the structure obtained by this method corresponds to those obtained by other methods.

The Perceptual Organization of Sound

Albert S. Bregman
Visiting Scholar, CCRMA
Dept. of Psychology, McGill University

I have been spending a sabbatical period at CCRMA, from September 1, 1993 to June 31, 1994. My main project is the creation of audio demonstrations of principles of human auditory perception to accompany my book on the perceptual organization of sound, *Auditory Scene Analysis*. They are intended to be used in the teaching of auditory perception at universities. They will be prepared in the form of a "mixed-mode" CD-ROM disk. One part of this disk will be playable as audio on any CD player. It will be synthesized on a PC compatible computer running William Henke's MITSYN software. A second part will be a HyperCard stack allowing interactive experimentation with audio using a Macintosh computer. The audio components for this part will be transferred to a Macintosh Quadra 950, where the HyperCard programming will be done.

The demonstrations relate to my research of the past 25 years on auditory scene analysis. Normally humans listen to sound in a background of other sounds. The spectrogram of the complex acoustic mixture that we encounter is like a picture that is created by taking the spectrogram of each of the individual sounds, drawing each on a separate piece of plastic, then superimposing them. To recognize the individual sounds, the human auditory system has to carry out a sort of decomposition of the incoming spectrum so as to recover the separate spectra that compose it.

Laboratory research using simplified signals has uncovered many of the principles by which the overall auditory data in partitioned into subsets representing separate environmental sounds. These principles can be illustrated by simple patterns of tones, each ten to twenty seconds in duration, formed of tones, glides and noise bursts that group or segregate perceptually for the listener, their organization changing as important acoustic parameters are modified. Each demonstration will demonstrate one or more principles of perceptual organization.

The demonstrations will be created in two formats, both encoded on the same "mixed mode" CD. Track 2 and onward will consist of about 40 demonstrations that are playable on any CD player. They will be explained in an accompanying booklet. The standard CD player will skip track 1. But when the CD is inserted in a CD- ROM player under the control of a Macintosh computer, an icon will come up on the user's screen and allow him or her to activate a HyperCard program stored on track 1, which will allow users to experiment with perceptual grouping. There will be about 15 demonstrations, similar to those on the plain-CD tracks, except that the user will be able to change a number of acoustic parameters in each demonstration and discover the perceptual results. Listeners will click on screen "buttons" to (a) start and stop the sound, (b) change parameters of the demonstrations, (c) read their descriptions, (d) learn relevant concepts, and (e) write and print notes that they make on their perceptions of the various sounds.

The acoustic material for both types of demonstration are being digitally synthesized on a PC-compatible computer running MITSYN software. This system allows users to create a network of boxes joined by lines on the display. The boxes represent simple acoustic instruments, and the lines represent the flow of "sound" from one to the other. It is a graphical implementation of the "unit generator" idea pioneered by Max Mathews. MITSYN also has a scripting language which allows all forms of signal processing to be implemented. It also has graphical waveform and vector editors.

The demonstrations intended for the fixed demonstrations will be played out of DAC's from the PC and recorded on DAT tape. They will then be mastered by a commercial company.

The ones destined for the Macintosh demonstrations will be transferred as disk files from the PC to the Macintosh and converted to sound resources. The set of demonstrations are being written in the form of a HyperCard stack. Special X- commands have been written which allow HyperCard to control looping and glitch-free concatenation of signals. Real-time parameter changes are accomplished as follows: volume and speed are done computationally, whereas changes in other parameters will be accomplished by swapping pre-computed files on the fly. This set of demonstrations will be delivered in CD-ROM format to the mastering studio.

Other projects that I am continuing involve my ongoing research and professional work which include reducing the backlog of research studies that have been generated in my laboratory but not yet written up for publication. Two recent ones involve the role of sudden onsets in the perceptual clarity of individual sounds in a mixture. Another concerns the relation between bottom-up and top-down processes in speech processing by the brain.

New and Revised Psychoacoustics Textbooks

John R. Pierce, et al.
Emeritus Visiting Professer of Music (Research)

The lecturers in CCRMA's Music 151 course, "Psychophysics and Cognitive Psychology for Musicians," have undertaken to write up the lectures for publication in book form, with lots of illustrations and including sound examples on a compact disc. This is proceeding, with preliminay writeups by Max Mathews, John Pierce, Roger Shepard, Perry Cook and others. We believe that this will provide an introduction and guide to psychophysics and cognitive psychology of musical sound and related matters that has not been available in the past.

In addition, John Pierce has revised his book *The Science of Musical Sound* originally published by Scientific American Library, and it has been published in paperback form by W. H. Freeman in 1992.

ARCHIVES

The Catgut Acoustical Society Research Library

Max V. Mathews
Professor of Music (Research)

Gary P. Scavone
PhD student, Music

Background and History

In 1990, the Catgut Acoustical Society began a search for a permanent home for its extensive files and research library. In 1992, the Center for Computer Research in Music and Acoustics (CCRMA) at Stanford University was selected for this purpose.

The Catgut Acoustical Society (CAS) was established in 1963 to assemble available knowledge and bring together researchers in string instrument acoustics. It is best known for its pioneering work in the research and application of scientific principles to the making of conventional and new instruments of the violin family. The organization's founder, Carleen Hutchins, is regarded as the world's foremost violin maker and as an expert on the acoustics of the instrument. She has personally made more than 400 instruments of the violin family.

The CAS archive makes up nearly 50 file drawers of information, including published papers, correspondences, meeting reports, extensive files on such people as Louis Condax, Robert Fryxell, and Felix Savart, a complete set of the CAS JOURNAL publications, and two Benchmark volumes of definitive papers in violin acoustics.

CCRMA has been working to supplement the CAS files with the personal archives of Arthur Benade and John Backus. Benade was a physicist working at Case Institute of Technology and Backus a physicist working at the University of Southern California. Both were world leaders in the acoustics of wind instruments. The addition of their archives would make the Stanford collection preeminent in the world for acoustical instruments. At the present time, CCRMA has verbal agreements for the transfer of Benade's and Backus' files, though a definite time frame has yet to be established.

Purpose and Goals

The Stanford collection has been established for the purpose of preserving and maintaining a complete and up-to-date repository of knowledge on musical acoustics. CCRMA is committed to honoring requests for information from interested parties all over the world, and further to incorporate new knowledge as it becomes available. It is intended that the archive be made as accessible as possible and its existence be made known around the world.

Activities

The process of transferring the CAS library to CCRMA and cataloguing its contents is currently underway. A computer database is being compiled and it is hoped that this file will be incorporated into an on-line computer network to allow for worldwide access.

The International Digital Electroacoustical Music Archive An Overview

Marcia Bauman
Research Associate

Max V. Mathews
Professor of Music (Research)

Background and History

The International Digital ElectroAcoustic Music Archive (IDEAMA) is dedicated to collecting, preserving and disseminating internationally renowned electroacoustic music. It was co-founded in December, 1990, by Stanford University's Center for Computer Research in Music and Acoustics (CCRMA), and the Center for Arts and Media Technology (ZKM), in Karlsruhe, Germany. The IDEAMA's initial goal is to create a target collection of early electroacoustic music composed during the period of approximately 1940-1965. Much of this music now resides on deteriorating analog tape. In an effort to preserve this music, it will be transferred to digital storage media, using CD standards. Information about the music will be entered into a catalog database. Although a large number of important electroacoustic works are currently accessible on analog tapes at the centers where they were produced, access to them is limited. The same is true for important works which are available through the private collections of individuals. Thus, one of the main challenges in developing the IDEAMA target collection is to identify and locate the desired materials so that an international collection of important electroacoustic music repertoire can be accessible at one single location.

The IDEAMA is unique in that it is a purely digital archive. All of the information, including sound, text and image, will be stored and retrieved in digital format only.

Organization

Founding Institutions

IDEAMA's administrative structure has been established to include three types of institutions: Founding institutions, Partner and Affiliate branch institutions. The founding institutions have collaborated in all the details of establishing policy and procedures for creating the archive and its ongoing function. CCRMA and ZKM are jointly responsible for collecting archive materials on a regional basis. ZKM focuses on European electroacoustic music, while CCRMA is responsible for music from the Americas, Asia and Australia.

Both founding institutions will provide each other with the materials they collect so that the complete target collection will be available at each location. Stanford University will house the IDEAMA at the Braun Music Center Archive of Recorded Sound. In Karlsruhe, the IDEAMA will be the core of the music section at ZKM's Mediathek.

Boards

To identify, locate and choose materials for inclusion in the target collection, each founding institution has formed a selection committee comprised of eminent composers, musicologists and other individuals who are well-versed and active in the field. In addition, internationally renowned composers and researchers form an overall international advisory board which establishes the international scope and reputation of the archive.

IDEAMA Branches

A partner institution has a collaborative relationship with the founding institutions, contributing materials to the target collection and/or participating in research for the catalog database. Partner institutions will also house the target collection and the catalog. After the target collection has been established and the catalog database implemented at the founding and partner institutions, an organization may become an affiliate branch by housing the target collection and integrating the archive database into its existing format.

Presently, there are seven formally designated partner institutions: the New York Public Library (NYPL); the National Center for Science Information Systems (NACSIS), in Tokyo; the Institut de Recherche et de Coordination Acoustique/Musique (IRCAM) in Paris; the Groupe de Recherches Musicales de l'Institut National de l'Audiovisuel (INA/GRM); the Groupe de Musique Experimentale de Bourges (GMEB); EMS in Sweden; and the Instituut voor Psychoacustica en Elektronische Muziek (IPEM) at the University of Ghent, Belgium.

The NYPL is planning an overall electroacoustic music collection. IDEAMA activities will take place within this context. Initial efforts will be made to digitize the works of such composers as Paul Lansky, Pauline Oliveros and Charles Dodge. NYPL will contribute to IDEAMA these and other works more readily available to NYPL on the East Coast.

In Japan, NACSIS serves as the nucleus for the nationwide comprehensive Science Information System, which covers the natural and social sciences, and the humanities. NACSIS links university libraries, computer centers, information processing centers and national university research institutions via computer and telecommunication networks. Research and development is carried out to compile databases and to create information systems.The NACSIS collaboration involves acquisition of electroacoustic music by Japanese composers

INA/GRM, formerly the Groupe de Musique Concrte at Radio-diffusion-Tlvision Franaise, was the first major location in the world for electroacoustic music production. Approximately 125 European works will be provided by INA/GRM.

IRCAM is one of the four components of the Centre National d'Art de Culture Georges Pompidou, together with the National Museum of Modern Art, the Public Information Library, and the Industrial Design Centre. It is an internationally renowned center for interdisciplinary research in the areas of computer technology, acoustics, digital sound synthesis, real-time digital signal processing, psychoacoustics, and computer music composition. In addition, IRCAM produces concerts, and its pedagogy department offers a series of educational symposia, seminars, courses and workshops for the public. IRCAM will provide access to European electroacoustic music and house the IDEAMA target collection and database.

GMEB was founded in 1972. Over the past 20 years, GMEB has sponsored an internationally acclaimed electroacoustic music festival and competition. Prize-winning works from GMEB have been broadcast in the "Le Chant du Monde" series and performed all over the world. Both IRCAM and GMEB will provide access to more recently composed European electroacoustic music.

The Target Collection

The original analog tapes for targeted works exist in a number of libraries, archives, radio stations, studios and private collections. Each founding institution has formulated a list of works based on their availability through these sources and upon the recommendation of each institution's selection committee. Approximately 800 works are presently being sought.

ZKM Selections

The European selection committee held its first official meeting at ZKM in May, 1992. At that time, committee members collectively recommended approximately 400 European works composed between 1930 and 1970. The two main criteria for selecting a work were its historical significance, and the urgency precipitated by the deterioration of the original analog tape on which it resides. Sources for the European works include major centers such as INA/GRM; Westdeutschen Rundfunks, Köln (WDR); and the Studio di Fonologia Musicale presso la sede RAI-TV, Milan.

In addition, works from smaller private and studio collections, and from the estates of Hermann Heiss and Joerg Mager will be included. The committee also decided to include, where possible, multiple versions of selected works and sound source materials, and later, film music, experimental radio plays and multi-media works.

CCRMA Selections

Several major centers have been contacted and arrangements are being made to digitize approximately 400 works and auxiliary materials. These centers include the Mills College Tape Music Center and the University of California, San Diego, which hold many of the pieces produced at the San Francisco Tape Music Center. Other centers include Columbia University's historic Electronic Music Center and the University of Illinois Experimental Music Studios. A number of significant works by Canadian composers are available through the National Library of Canada. The Laboratoris de Investigacion y Produccion Musical (LIPM), the first major center for Latin American electroacoustic music, has digitized approximately 30 works for the target collection.

The personal collections of Max Mathews and Gordon Mumma provide a wealth of historically significant electroacoustic music for the IDEAMA. Max Mathews has contributed tapes of the earliest computer sound and music developed at Bell Laboratories. Gordon Mumma's collection includes taped performances by the ONCE Theater Group, from Ann Arbor, Michigan, which was a major performance group in the '60s "happenings" scene. CCRMA will contribute other early computer generated works. Works by composers who were not necessarily associated with major centers, but whose contributions are no less significant, will also be included.

NACSIS is researching and acquiring Japanese works, while research at CCRMA has been initiated to identify Australian works for the target collection. Works by Australian composer Tristram Cary have been digitized and arrangements are underway to include works of other Australian composers.

A commercial multi-channel digital recorder will be used for the preservation and playback of multi-channel works. In addition, the multi-channel works will be stored, not as sound files, but as data files, on CD-ROM, which can be transferred to computer disk. At CCRMA we have recently developed the capability to play four channels using the NeXT computer.

Future Plans

The IDEAMA is envisioned as an archive that will be a major research tool for the international scholarly community, and one which has the widest possible access. Our present plans at CCRMA include: 1) transferring the target collection to recordable CD's and/or CD-ROM disks; 2) providing at least three copies of disks (to Stanford University's Archive of Recorded Sound, ZKM and the New York Public Library); 3) creating a simple computer database for the collection; 4) permanently housing the collection at the Stanford University Archive of Recorded Sound. The Archive of Recorded Sound will have a computer terminal to allow automated access to the collection and equipment to make exact digital copies of the works in the collection.

The IDEAMA originally intended to create a catalog in MARC format and to computer-scan auxiliary paper materials (program notes, scores, etc.) and storing them on CD-ROM. Due to lack of funding, these objectives have been eliminated or postponed. Copies of paper adjuncts to the sound will be kept as traditional library documents for possible future scanning. In the meantime, IDEAMA materials will be a testbed for a multi-media distributed library project currently being developed at Stanford University's Academic Computing Center. It is our hope that once the target collection is completed, more recently composed works will be added in order to represent the field from its inception to the present day.

Acknowledgements

At Stanford University, the IDEAMA is supported by CCRMA, the Andrew W. Mellon Foundation and the National Endowment for the Arts. ZKM receives its support from the state of Baden-Wuerttemberg and the city of Karlsruhe.