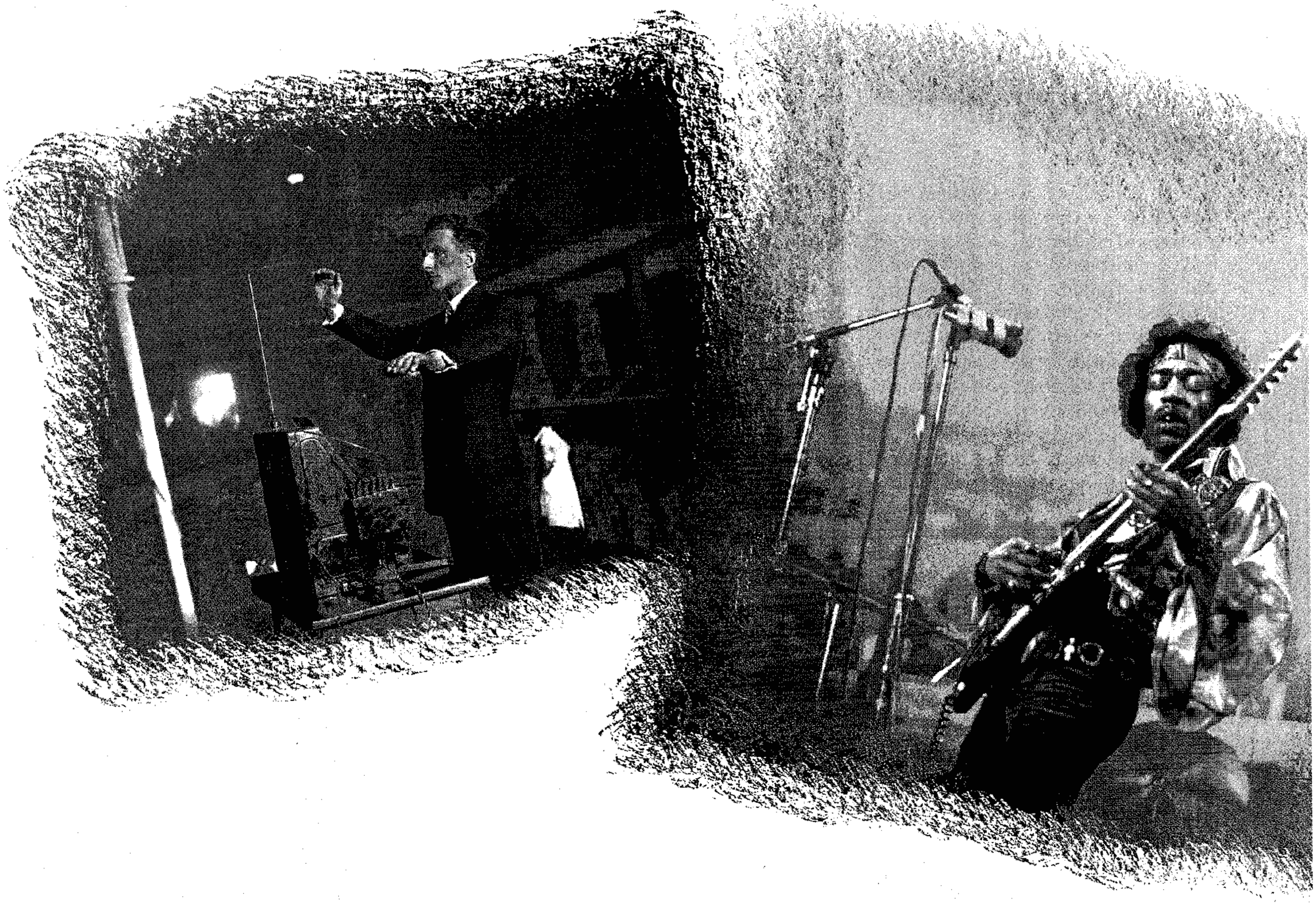


Electronic music:

The vast soundscape that synthesizers produce now can be explored with the full range of human gesture

cover story

Surely one of the earliest applications of toolmaking involved making music—controlling sound of one sort or another. Whenever this quest started, it still thrives, as researchers and developers rapidly apply new technologies to musical performance.

The most common form factors (the ergonomics, as it were) of electronic music controllers grew out of those in standard instruments, such as keyboards, guitars, and wind instruments. But new technology has freed instrumentalists from the parameters of standard instruments. With new controllers people can play synthesizers—produce any sound event they want—by leaping through the air, swinging their arms, or even buttoning up a Levi's jacket.

Modern electronic instruments, namely, music synthesizers and computers, essentially allow composers and performers to create *any* sound or sound

0018-9235/97/$10.00©1997 IEEE

[1] What's past is prolog in revolutionary music controllers. Back in the 1920s, Leon Theremin created one of the most radical electronic instruments ever, with pitch and amplitude that were controlled by moving the hands through space over sensor plates and oscillators. [far left, with inventor].

Forty years later, body and instrument join again in a driven system, when Jimi Hendrix alters his guitar sound by moving his body to generate feedback.

Today, free-gesture control returns with the Sensor Chair synthesizer controller, where the body *is* the transmitter in a closed-loop electrical field.

PHOTO CREDITS:
CORBIS/BETTMAN; RETNA LTD.;
JOHANNES KROEMER.
DIGITAL MANIPULATION:
BARKING LION, INC.

JOSEPH A. PARADISO
Media Laboratory
Massachusetts Institute of Technology

sequence. But how do they interact with the core of these instruments, the actual sound-production circuitry, especially if the sounds are to be controlled in live (real-time) performance? Enter the interface.

"Interface," in a more conventional sense, might comprise both Microsoft Windows and a mouse. Most mice, on their own, are limited controllers, capable of doing just a few things. That is, they have few degrees of freedom: they map two-dimensional cursor location and, with a click, a "do it" command, all done with few physical gestures.

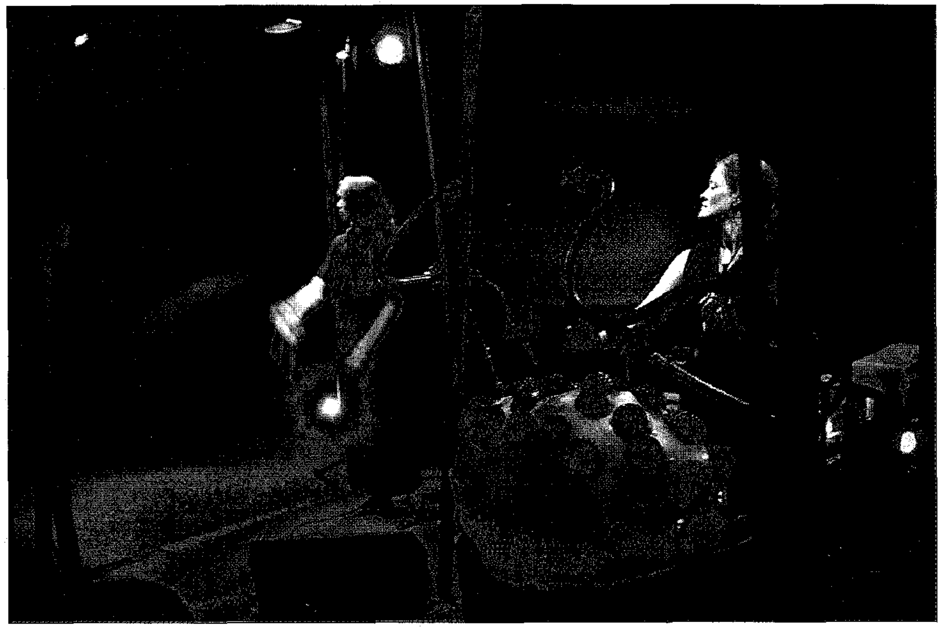
But a musical instrument usually has many more degrees of freedom. Violinists can interact with their sound-producing equipment (strings) using a broad range of physical gestures and choosing from a huge range of bowing pressures, pitch choices, and so on. Many musicians already have well-developed techniques for playing acoustic instruments—they have mastered minute control of many parameters.

As the devices to be discussed below reveal, those parameters, conveyed either in the acoustic sound produced or in the physical gestures themselves, can be digitally captured and sent on to the synthesizer. There the parameters can be mapped to whatever the performer desires: a percussionist plays a drum roll and hears a rippling cascade of notes, each impact mapped to a pitch; an acoustic cello stroke

Web *Enhanced*

Turn to the on-line version of this article at http://www.spectrum.ieee.org and check out over 20 video and sound clips—as well as two different ways to interactively make music over the Web.

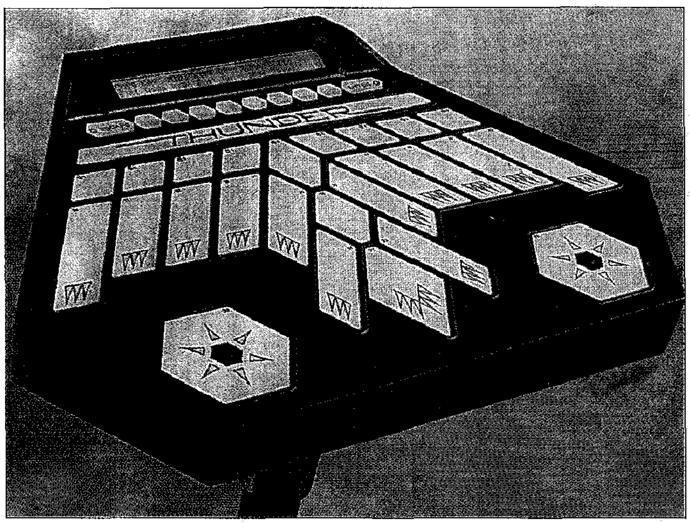
new ways to play



[2] The MIT Media Lab's *Brain Opera* swings into gear with the lab's custom music controllers. In the foreground is the 300-pad Rhythm Tree percussion controller. Above this instrument is a Gesture Wall, which the percussionist also plays by moving her hands through the space between its flower-like receivers. At the rear a conductor guides the entire ensemble with a Digital Baton.

JOHANNES KROEMER

[3] In Donald Buchla's Thunder percussion controller [below], different areas (zones) respond separately to strike velocity, location, and pressure.



sets off a synthesized orchestral crash; or a performer waves her hands downward in free space and a sound gets softer.

Synthesizers communicate with controllers (and with other sound-producing electronics) via the Musical Instrument Digital Interface standard (MIDI). This milestone standard in electronic music, introduced in 1983, enables equipment to communicate across a daisy-chained serial bus. MIDI messages essentially are parameters that define musical events and tell each synthesizer what kind of sound to create: one message could be to turn on a sound, another to change its spectral content, and another to turn it off.

Of the many ways in which MIDI has revolutionized electronic music, perhaps the most relevant here is how simple it makes including a computer in the loop between human physical action and musical response. With this the definition of musical performance changes: even the simplest gestures of musicians without classical training can be translated into wonderfully complex musical events. By the same token, the kinesthetic control of virtuosos can be extended to manipulate entire synthesized soundscapes. Early electronic music pioneers glimpsed this, and we can gain perspective today by understanding their quest for ever-new ways of controlling electronic sound expressively.

Reign of the keyboard

With the exception of the Theremin [Fig. 1], to be discussed below, early electronic musical instruments were primarily controlled by keyboards—often, the standard layout familiar from acoustic pianos.

The first true electronic instruments were Elisha Gray's 1876 Musical Telegraph and William Duddel's 1899 Singing Arc. The Telegraph was an array of tuned electromechanical reeds connected to a keyboard. Duddel's invention imposed an audio modulation frequency over a 300-V potential driving a carbon arc lamp to produce musical tones.

In 1906, Thaddeus Cahill brought forth the 200-ton Telharmonium. It generated musical audio from a building with 145 tuned dynamo wheels, whose shafts and inductors produced current at audio frequencies. Cahill's was the first electronic instrument with a velocity-sensitive keyboard, through which a player was able to express changes in loudness (dynamics). The feature would not be seen in popular instruments for decades.

The first electronic instrument to be produced in any quantity was the Ondes Martenot, introduced in 1928 by Maurice Martenot. Notable composers used it in their works, and it is still used to perform them [AUDIO 1]. The instrument's arrangement prefigured that of almost every keyboard synthesizer today: right hand for basic frequency (pitch), and left hand for providing articulation features that change the character of the sound.

Early versions controlled monophonic pitch (that is, only one note at a time) through a taut ribbon loop attached to a ring on the performer's right forefinger and wound around the shaft of a variable capacitor—much like the string in the tuning dials of old radios. The left hand flipped switches controlling amplitude dynamics and timbre, the combination of harmonics that make up the distinctive "sound" of a pitch beyond that given by a pure sine wave.

Keyboards continued to advance, with one of the more interesting appearing in George Jenny's Ondioline, a popular 1941 French instrument. With net key pressure the entire Ondioline keyboard moved up and down, and with lateral force it moved side to side. The capacitance was thereby varied between sets of sensor plates, thus allowing dynamics to be mapped onto key pressure and vibrato effects to be determined by horizontal displacement (vibrato is the musical term for a minute warbling in pitch).

No summary of milestones in musical interface technology would be complete without mention of Hugh LeCaine's 1948 Electronic Sackbut. The instrument gave access to electronic sound with position- and pressure-sensitive keys, a horizontally movable keyboard to play new sets of tones, foot pedals, and a novel capacitive touch controller for the fingers of the left hand. With all these continuous input channels modifying its timbre and dynamics, the Sackbut, as LeCaine's recordings show, came alive in the hands of talented performers [AUDIO 2].

For the synthesizers of the late 1960s and early '70s, the dominant interface was once more a simple piano-like keyboard. Although a few of the keyboards could transmit requests for two notes at once, most made do with one, and most lacked any response to velocity, pressure, or indeed anything besides key hits.

When a key on one of these devices is struck, it generates a gate pulse while grabbing a voltage off a resistor ladder with a sample-hold. The gate triggers envelop generators that control the amplitudes; the keyboard's sampled voltage determines pitch, filter tracking, and so on.

Continuous control of timbre (the overlays of spectra) was achieved by turning knobs or manipulating something resembling a ribbon controller, a resistive strip that responds to finger position. In a few of these synthesizers, such as those built by Donald Buchla, a near-legendary designer of novel electronic interfaces, keyboards were replaced entirely by flat pressure- and position-sensitive capacitive touch-plates.

To change their musical characteristics, early analog electronic "modular" synthesizers depended on patch cords connecting large banks of equipment. Responding to the demands of touring bands for a portable synthesizer, in 1970 Moog Music Inc., Cincinnati, Ohio, released the MiniMoog, a hard-wired subset of earlier modular systems. Trading portability for flexibility and sonic variety, the MiniMoog dictated signal routing between components by means of switches and potentiometers.

It was such a success that it has cast design shadows to the present day. The most obvious is the twin set of wheels to the left of nearly all electronic-music keyboards. Ordinarily, one wheel controls continuous changes in pitch (known as pitch bend, when the sound "slides" through frequencies); the other controls oscillators and filters to change the tone quality. (Current synthesizers can remap these wheel controllers to any desired function.)

In the early '70s, synthesizer keyboards finally started allowing players to produce several pitches at a time. This "polyphonic breakthrough" was made possible by Dave Rossum and Scott Wedge of E-Mu Systems, Scotts Valley, Calif., and grew out of their work with digital scanning keyboard technology pioneered by Ralph Deutsch at the Allen Organ Co., Macungie, Pa.

As for dynamics in a modern keyboard, under the MIDI standard all detected note events are accompanied by a 7-bit velocity parameter. The velocity—which amounts to an indicator of "how hard" the key is hit—is measured by clocking the amount of time it takes for a key to switch between an upper and lower contact. Besides indirectly measuring the force on the key, it monitors when the finger is removed.

Very little hardware is needed to measure and transmit the velocity parameter. As a result, commercial synthesizers finally could provide some sort of dynamic response in a way familiar to pianists (unlike the ubiquitous electric organs). Many keyboards also send other parameters supported by the MIDI standard, including "aftertouch," which is the pressure on the key after its initial strike, a parameter not present in pianos or organs.

Much research is under way to make keyboards more expressive. For example, the Multiply Touch Sensitive Keyboard, by Robert Moog and Thomas Rhea (who is now at the Berklee School of Music, in Boston) capacitively measures the position of the finger in the *x* and *y* directions on top of each key. And such researchers as Brent Gillespie at the Stanford University Center for Research in Music and Acoustics (CCRMA) are exploring the use of electromechanical servos to design keyboards with programmable mechanical responses, so as to change or improve the "feel" of the keyboard to the player.

Finally, several groups are changing the layout of keyboards entirely—for instance, adding more keys to make playing microtonal music easier (microtonal music divides the octave into more pitches than the 12 tones common in the West). In the SalMar Construction, designed in 1969 by the late Sal Matirano at the University of Illinois, the sound is manipulated by a bank of 291 switches through which the performer controls 24 channels of synthesized audio—an extreme adaptation of a "keyboard" [VIDEO 1].

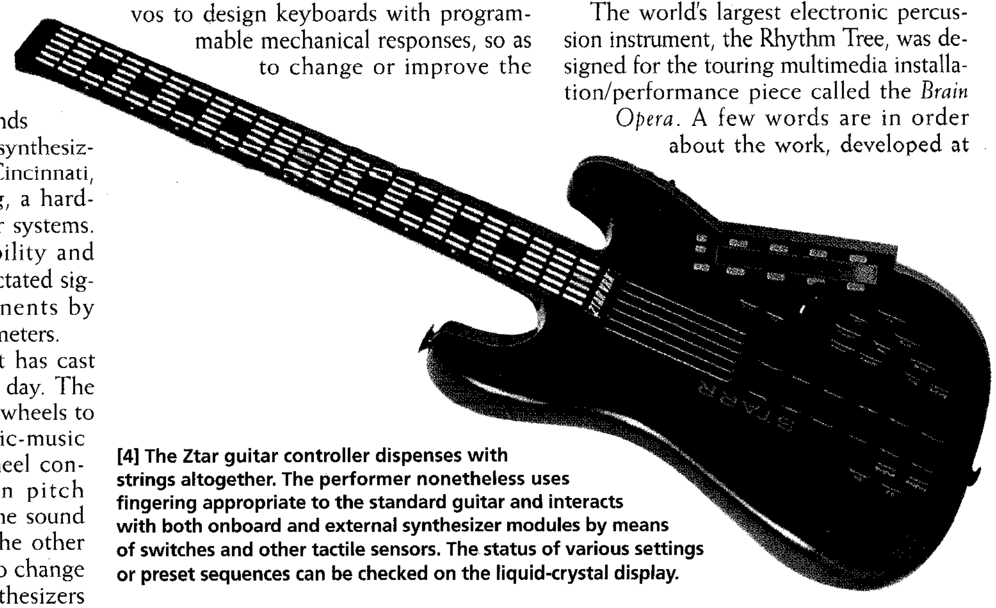
Percussion interfaces

Non-keyboard percussion instruments began to sprout electronic enhancements by the late '60s. Percussionists started triggering electronic sounds by way of acoustic pickups on the surfaces they struck; the trigger pulse, as well as voltages proportional to the strike intensity, were routed to synthesizer modules to produce sounds.

Electronic percussion took a major leap in the early '80s with the designs of Dave Simmons, who embedded synthesizers with innovative new sounds into a drum controller. The player could use traditional drum techniques on its flat, elastic drumpads. But any actual sounds emanating from the pad were acoustically damped. Rather, the electronic signals transducted from the pads drove the synthesized audio, chosen at the player's whim. Simmons soon added the ability to send MIDI output to external synthesizers.

Later versions introduced the concept of zoning: hits of varying intensity in different areas of a single pad trigger different sonic events chosen by the drummer. Today, nearly every musical instrument manufacturer has a line of electronic percussion interfaces. Most use force-sensitive resistors as sensing elements; some incorporate piezoelectric pickups.

The world's largest electronic percussion instrument, the Rhythm Tree, was designed for the touring multimedia installation/performance piece called the *Brain Opera*. A few words are in order about the work, developed at



[4] The Ztar guitar controller dispenses with strings altogether. The performer nonetheless uses fingering appropriate to the standard guitar and interacts with both onboard and external synthesizer modules by means of switches and other tactile sensors. The status of various settings or preset sequences can be checked on the liquid-crystal display.

the Massachusetts Institute of Technology (MIT) Media Lab under Tod Machover's direction. For one thing, it uses many of the instruments described here. Some are played in a free-for-all setting by the public, and others in a concert setting by musicians trained on them [Fig 2].

At times, music recorded from the on-site, "naive" walk-ins was combined with the live performance of the trained musicians, music by Machover, and music submitted to the ensemble created on-line by logged-in users of the World Wide Web. The entire performance was then transmitted live on the Internet [see "The Web resounding," p. 27].

The *Brain Opera*'s percussion instrument, the all-digital Rhythm Tree, features an array of more than 300 smart drumpads [Fig. 2, foreground/VIDEO 2]. Each drum pad has an 8-bit microcontroller that analyzes the signal from a piezoelectric foil pickup and drives a large light-emitting diode (LED), all potted in a translucent urethane mold.

All the pads, which are mounted on a daisy-chained RS-485 bus, are sequentially queried with a fast poll. After performers smack one of the pads with their hands, the pad responds with the degree of force used and the zone of the pad that has been struck. A MIDI stream with these and many other parameters is then produced, which triggers sounds and flashes lights in the pads for aesthetic visual feedback.

A somewhat more complex—and thus expressive—contemporary percussion controller is Donald Buchla's Thunder [Fig. 3]. Unlike the thick, nodule-like pads on the Rhythm Tree, Thunder has hand-sized zones that respond separately to strike velocity, location, and pressure. The latest versions use electro-optical sensing to detect the surface membrane's deformation under hand contact. Thunder gives the player many ways to assign (map) many kinds of complicated percussive events into a MIDI data stream, which is then interpreted by the synthesizer.

No strings attached

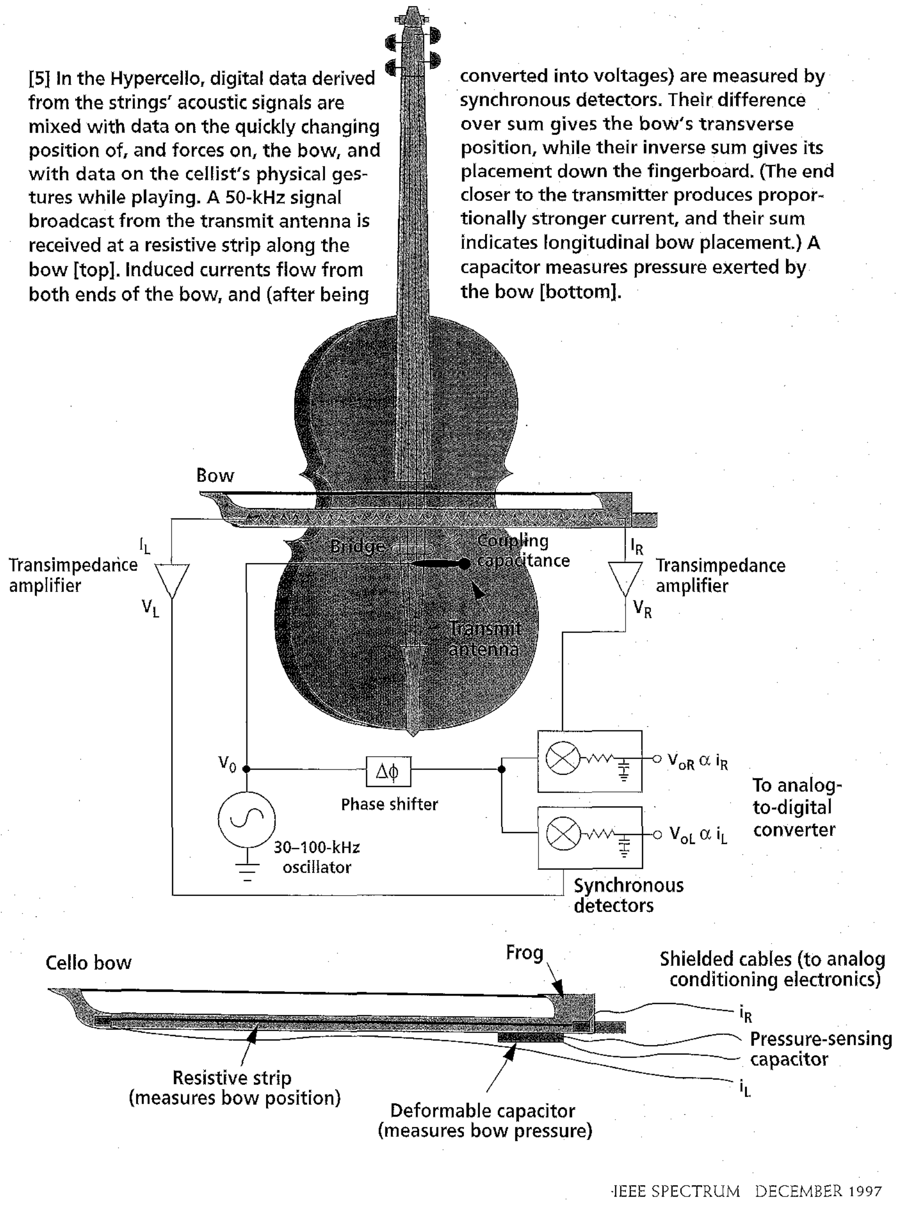
Stringed instruments have followed a technically challenging road into the world of electronic music controllers. The world saw the birth of an important and in time wildly popular electronic instrument when guitars were mated to electronic pickups back in the '30s. The instrument was an extreme break with tradition, and its musical possibilities were pioneered by innovative musicians such as Charlie Christian and Les Paul. But many sounds remained latent in this new instrument, so to speak, until the '60s, when guitarists started exploring the virtues of distortion and feedback—with Jimi Hendrix recognized as the early master of these techniques [Fig. 1, center]. In this performance style, the electric guitar becomes part of a complex driven system: the instrument's behavior depends on factors such as the room acoustics and the location of the players themselves, particularly their position relative to the loudspeakers.

A wealth of analog gadgets and pedals came on the scene to modify the signal from the guitar pickup. Among them were wah-wahs (sweeping bandpass filters), fuzzboxes (nonlinear waveshaping and limiting), and flangers (analog delays or comb filters).

Still, guitarists demanded access to the sonic worlds opened up by synthesizers. The most important step toward this goal was extracting the many parameters of the vibrating guitar string: the frequency and amplitude, obviously, both steady state and over time, as well as distinctive features defining the initial pluck or pick of the string (the attack). This is still a challenge to do quickly, accurately, and cheaply.

The problems arise at several levels, including noise transients with the attack of the sound, the potential need for several cycles of a steady-state waveform for robust pitch determination, dealing with variabilities in playing style, and the difficulty of separating the sounds from the different strings.

Rough attempts in the early '70s were technical disasters, of very limited musical applicability. But in the later '70s a remedy was introduced that is still in use today: the hexaphonic magnetic pickup. The pickup has one coil for each string, thus producing six independent analog outputs, and is mounted very close to the strings and bridge to avoid crosstalk. Some other pickup designs use optical or piezoelectric sensors to track string vibration. These two sensory modes can provide more robust signals and work with nylon or other nonmetallic strings.



[5] In the Hypercello, digital data derived from the strings' acoustic signals are mixed with data on the quickly changing position of, and forces on, the bow, and with data on the cellist's physical gestures while playing. A 50-kHz signal broadcast from the transmit antenna is received at a resistive strip along the bow [top]. Induced currents flow from both ends of the bow, and (after being converted into voltages) are measured by synchronous detectors. Their difference over sum gives the bow's transverse position, while their inverse sum gives its placement down the fingerboard. (The end closer to the transmitter produces proportionally stronger current, and their sum indicates longitudinal bow placement.) A capacitor measures pressure exerted by the bow [bottom].

Because of the limited signal-processing technology available at the time, these interfaces remained too sluggish and inaccurate for many guitarists. As a result, by the mid-'80s, guitar controllers entirely bypassed the problem of detecting pitch. Instead, the player's left-hand finger positions on the fretboard were sensed typically by way of switches or capacitive strips under the fretboard. A pickup was still used, but now only to measure the amplitude of the string's vibration. Some ingenious designs tracked the fret fingering by timing ultrasound pulses propagating down the string from the bridge and reflecting back where the string is pressed against the fretboard.

The most extreme forms of these controllers disposed with strings entirely. Only the guitarist's gestures are measured, through switches and tactile sensors. Although many manufactures made these interfaces, almost none are still in production, a notable exception being the Ztar series from Starr Labs, San Diego, Calif. [Fig. 4/VIDEO 3].

In addition, advances in digital signal processing technology and pitch extraction algorithms have significantly improved the performance of today's pickup-based interfaces. These can be retrofit to standard guitars, providing a response that is much faster and more robust than their predecessors. Although not quite all guitarists are satisfied, many find these interfaces playable and useful.

It is generally accepted, however, that the MIDI standard cannot handle the wealth of data that stringed instruments, including guitars, can produce. A guitar performance can be more or less shoehorned into a set of real-time features (say, notes and attack amplitudes) that fit the MIDI standard. But it does not provide enough bandwidth to transmit the numerous channels of detailed, continuous articulation these instruments can produce.

Several solutions have been suggested, such as the currently dormant ZIPI interface standard, proposed several years ago by the Center for New Music and Audio Technologies (CNMAT) at the University of California, Berkeley, and Zeta Music Systems, Oakland, Calif. Unfortunately, none has yet been accepted, and the world of high-bandwidth musical controllers still awaits a successor to MIDI.

The classics updated

Players of orchestral instruments and synthesizers inhabited separate domains during their early courtship. If synthesized and traditionally generated music were used in a composition, the live musicians essentially kept time to a tape recording of the generated music or to a prerecorded musical sequence played on the fly by the synthesizer.

This relationship has now become a good deal more intimate. The best-known large-scale examples of computer/human ensemble integration were provided by Giuseppe Di Giugno's 4X synthesizers, which were developed in the '80s in Paris at the Institut de Recherche et Coordination Acoustique/Musique (IRCAM). Pierre Boulez was one of the many composers who used the 4X to analyze and process audio from acoustic ensembles, enabling musicians to control real-time synthesized contributions.

That trend is continuing, as more flexible interfaces between classic orchestral instruments and computers emerge. With new controllers, synthesizers can now be a virtual accompanist and partner, able to adapt and respond to the detailed nuances of individual musicians.

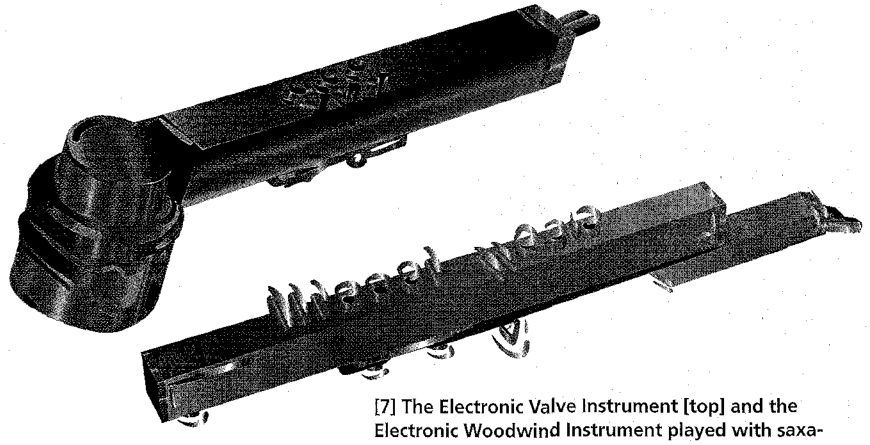
Early electronic interfaces for bowed string instruments processed only their sound, adding amplification and effects. Granted, the complicated and dynamic nature of a bowed sound makes fast and robust pitch tracking difficult. All the same, many researchers have developed software with this aim. Commercial MIDI pitch-tracking retrofits are also manufactured (for example, by Zeta) for violins, violas, and cellos.

Nonetheless, signal processing alone is inadequate for prompt and full translation of gesture in string bowing. The string controller must capture as many bowing parameters as possible if the synthesized sound is to respond promptly to the full range of the player's musical nuances—for example, those performed by selective direction and placement of the stroke, applied pressure, and angle of attack.

Rather than trying to infer the bowing dynamics directly from the audio stream, researchers have developed sensor systems to directly measure the bowing properties. A design by Chris Chafe, of CCRMA, for one, employs accelerometers and an infrared tracker to measure the motion and position of cello bows. Peter Beyls, at St. Lukas Art Institute, Brussels, has built a violin using an array of infrared transmitters and receivers to detect bow and fingering gestures.

Researchers at the MIT Media Lab have used several technologies to design bowed-instrument controllers. These efforts began with the Hypercello, designed in 1991 by Neil Gershenfeld and his colleagues. Audio signals from each string

[6] The Hyperviolin, being played here by Ani Kavafian, operates like the Hypercello, apart from an untethered bow. Three transmitters are used. Two on the bow, transmitting at different frequencies, are used to gauge its location. Bow pressure is measured by a force-sensing resistor.



BARKING LION, INC.

[7] **The Electronic Valve Instrument [top] and the Electronic Woodwind Instrument played with saxaphone and trumpet technique, respectively, make no sounds on their own. Rather, the physical gestures are captured and interpreted by the synthesizer.**

are individually detected, and their sonic characteristics extracted, but so are a host of gestural parameters. Positions of the left-hand fingers are measured by a set of resistive strips atop the fingerboard, and the bow position and placement is revealed when a resistor strip within it is capacitively coupled to a transmitted signal from the instrument's bridge (this technique is also less prone to error than the accelerometers and infrared of other bow-sensing controllers) [Fig. 5].

In the Hypercello, a 50-kHz signal is broadcast from an antenna atop the bridge and received at a resistive strip running the length of the bow. Induced currents flow from either end of the bow. After being converted to voltages by the transimpedance amplifiers, they are measured by synchronous detectors, one for each side. Their difference-over-sum indicates the transverse bow position, because the end closer to the transmitter produces proportionately stronger current. Their inverse sum indicates the bow's placement up and down the fingerboard, since net capacitive coupling decreases as the bow moves down the instrument, away from the transmitter.

A deformable capacitor beneath the player's bowing finger measures the applied pressure (to gauge the dynamics), and an Exos Dexterous Wrist Master on the left wrist digitizes the angle of the player's wrist—an important parameter for cleanly extracting certain gestures such as vibrato. The Hypercello made its debut in 1991 at the Tanglewood Music Festival, Lenox, Mass., where cellist Yo-Yo Ma performed Tod Machover's composition from that year, *Begin Again Again...* [VIDEO 4].

The bow was originally tethered by cables to the signal-conditioning electronics. This was tolerable for cellists but awkward for violinists. A wireless bow tracker, also developed at the Media Lab, remedied the problem with three small battery-powered transmitters on the bow and a receive electrode on the violin, above the bridge. Two transmitters broadcast at different fixed frequencies, which drive either end of the resistive strip.

Using the same sum and difference analysis employed in the Hypercello bow, the amplitude of the components in the received signal indicates the bow's horizontal and vertical positions. A force-sensitive resistor placed below the player's bow grip causes the frequency of the third oscillator to vary with applied pressure, and an FM receiver tracks these changes. Violinist Ani Kavafian used this system in several performances of Machover's 1993 *Forever and Ever* [Fig. 6/VIDEO 5].

Interfaces for wind players

Electronic music researchers were quick to design controllers so that wind players could use their technique to control electronic sound. Wind instruments are essentially monophonic, and were well-suited to driving the single-voice synthesizers and the effects processors produced in the '60s and '70s. These systems were playable, but the early pitch trackers with which they were used were easily confused by harmonics, attack transients, and other performance artifacts.

Following the familiar pattern, by the early '70s designers found it easier to produce controllers to capture the player's gestures—the state of valves, holes, and the changing pressure of the player's breath—without the instruments generating any sound worth noticing. Valves no longer regulated the length of a vibrating air column, but became switches; mouthpieces and reeds became breath and bite sensors.

The first controller to come to public attention was the Lyricon Wind Synthesizer Driver. It produced voltages from changes in fingering, lip pressure, and air pressure, which was sensed through a hot-wire anemometer. Originally packaged solely as a controller, it was later mated with a small dedicated synthesizer.

In the late '70s and early '80s the trumpeter Nyle Steiner developed two of the best-known electronic wind controllers: the Electronic Woodwind Instrument, with a saxophone-like fingering, and the trumpet-like Electronic Valve Instrument [Fig. 7]. In addition to breath- and lip-pressure sensors, these instruments feature capacitive touch keys to make it easy to play fast; touch plates and levers help in playing *portamento* (letting notes "slide" more or less from one to another), vibrato, and other effects. Akai still produces a version of the Electronic Woodwind Instrument.

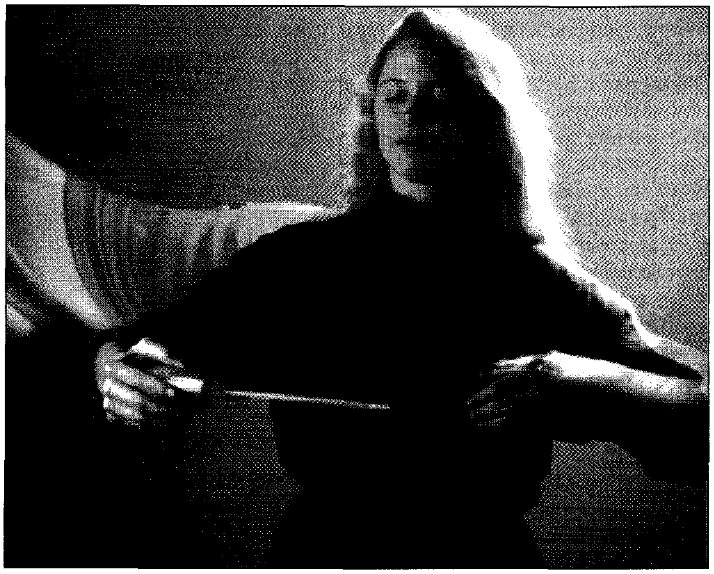
Yamaha played a considerable role in the emergence of digital wind interfaces, first by introducing simple breath controllers to provide an additional way to control sounds made by the company's popular FM keyboard synthesizers. In the late '80s, the company introduced the WX-7, the first wind controller to output MIDI, with a fingering layout close to a traditional saxophone. A later version, the WX-11, is still in production.

Taking up the baton

As synthesized sound has become more complicated, composition algorithms have grown better at specifying ever more refined musical details. This suggests great opportunities for interfaces like the familiar orchestra conductor. This interface—actually, the human being, who sometime uses a baton—dictates high-level structure. "High-level" changes might be transitions in sound groupings or underlying rhythm. The classical conductor implements these high-level musical commands through the timing, placement, sweep, and acceleration of beats. The "details" (the sounding tones) are handled by the instrumentalists—the synthesizer.

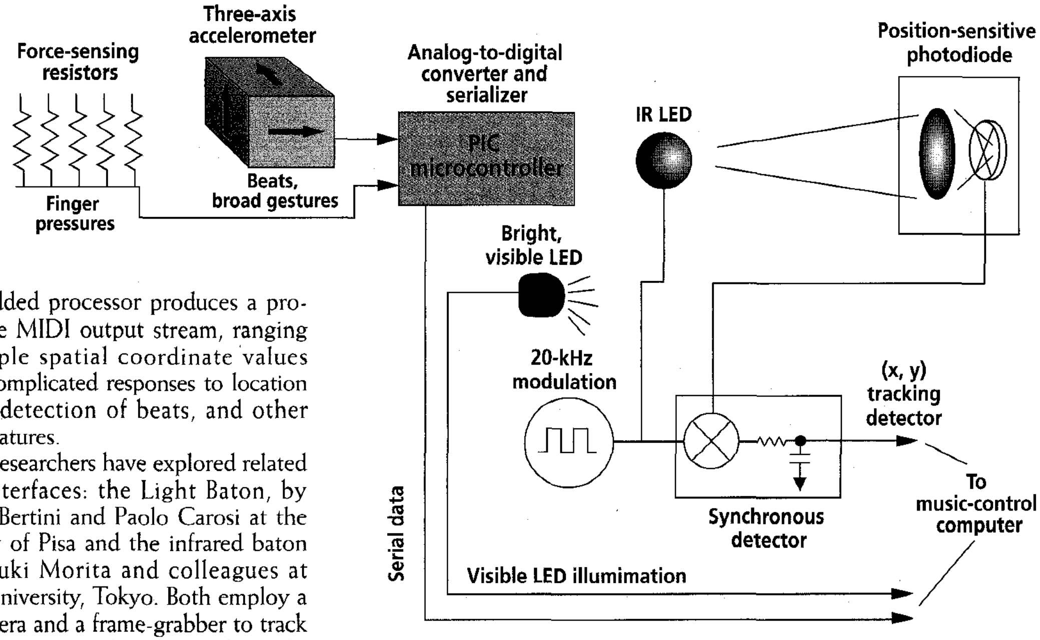
Perhaps the most famous electronic baton controller was developed in the late '80s by Max Mathews and Bob Boie, then at Bell Laboratories. Their Radio Baton system independently measures in three axes the continuously changing positions of a pair of wands. One is held in each hand above a sensing platform. The wands exploit capacitive proximity sensing: an array of receiver electrodes in the sensing platform detect the signals from transmitters (broadcasting at different frequencies) in the wand tips. By suitably processing the received signal strengths, the system can determine the position of the wands in real time.

Instead of electrical coupling, optical tracking is the basis of Donald Buchla's commercially produced Lightning system: a pair of wands whose vertical and horizontal positions can be continuously sensed across a performance area of 4 by 6 meters. Each wand has a modulated infrared-LED on the tip, which is tracked by a photodiode array in a base station.



[8] A "conductor" wielding a Digital Baton works with acceleration, light, and finger or hand pressure. The precise position of the baton tip in two-dimensional space is tracked by an infrared light-emitting diode [LED] and a photosensor. The visible LED on the tip both gives directional feedback to the conductor and brightens in response to pressure on the baton. Beats and large gestures in all directions are monitored with the accelerometers. Five force-sensing resistors encased in the baton's urethane skin react to pressure from individual fingers or the palm of the hand.

WEBB CHAPPELL



An embedded processor produces a programmable MIDI output stream, ranging from simple spatial coordinate values through complicated responses to location changes, detection of beats, and other gestural features.

Other researchers have explored related optical interfaces: the Light Baton, by Graziano Bertini and Paolo Carosi at the University of Pisa and the infrared baton of Hideyuki Morita and colleagues at Waseda University, Tokyo. Both employ a video camera and a frame-grabber to track the two-dimensional motion of a light source at a wand's tip. Like most software vision systems, these can exhibit marked processing delay and sensitivity to background illumination.

Instead of watching for beats with a spatial tracker, some batons use accelerometers to feel them. These systems include three-axis devices built by Hideyuki Sawada and his colleagues at Waseda University; the MIDI baton by David Keane and his colleagues at Queens University at Kingston, Ontario, Canada; and the Airdrum, a commercial dual-wand unit by Palmtree Instruments, La Jolla, Calif.

A belt-and-suspenders approach incorporating optical trackers and accelerometers is used in the Digital Baton, built at the MIT Media Lab for the *Brain Opera*. Using a synchronously demodulated position-sensitive photodiode, the system tracks the horizontal and vertical position of an infrared LED at the baton's tip [Fig. 8]. To detect large-scale gestures promptly, it uses a three-axis 5-g accelerometer array. Also, five force-sensitive resistors are potted in the urethane baton skin to measure continuous finger and hand pressure. With such a wealth of continually variable features, the device is a very versatile controller [VIDEO 6].

Holding nothing at all

Some controllers lack even the slightest physical resemblance to any tried and true acoustic forebears—a break with tradition that opens up fascinating channels for expression. Several types respond to body position and motion, without requiring anything to be grasped or worn. These so-called free-gesture controllers cannot be played with as much precision as the controllers described above—generically called tactile controllers—nor are they designed to be as precise. But with the aid of a computer interpreting the controller data and applying interesting sonic mappings to them, very complicated musical events can be launched and controlled.

Sensing technologies from capacitive sensing to machine vision have been turned to advantage in these controllers. And because many of the technologies and computer resources have become so inexpensive, they are becoming more common. For example, they often are used for interactive installations open to the general public—such as corporate lobbies in which office workers create music simply by walking to the elevators.

Indeed, the interface that foreshadowed much of the electronic revolution in music-making was a free-gesture controller: the Theremin, developed in the early '20s by the Russian physicist, cellist, and inventor Leon Theremin. Its player's hands move freely through the air above a pair of sensor plates, one hand controlling pitch and the other, amplitude [Fig. 1, again].

The hands capacitively couple into the plates, with the coupling decreasing as a function of the distance between hand and plate. As the capacitance at a plate changes, a connected resonant oscillator is detuned. One oscillator produces a heterodyned audio beat for pitch; the other, determining amplitude, uses the amount of detuning to gate the audio signal.

Manufactured by RCA Corp. a decade after its invention, the Theremin was an international sensation. It spawned several successful performance careers, most notably that of its primary virtuoso, Clara Rockmore, as well as Theremin himself [VIDEO 7]. (Played in large swoops of pitch, its sound became a cliché for spookiness in the science-fiction movies of the '50s.)

The Theremin still has an enthusiastic following. In fact, Robert Moog, the famous synthesizer pioneer, began his career in electronic music in the '50s by building Theremins, and his company, Big Briar, Asheville, N.C., still sells them.

Electrical fields forever

The MIT Media Lab has developed many musical interfaces that generalize Theremin's capacitive techniques into what can be called electric field sensing. The Theremin works through a loading mode: in essence, it detects current pulled from an electrode by a nearby capacitively coupled human body. The Lab has based its devices on transmit and shunt modes, which provide more sensitivity and longer range.

The transmit mode takes advantage of the human body's high conductivity. When someone sits or stands on a transmitter electrode, that person in essence becomes an extension of the transmit antenna. The signal strength induced at a receiver electrode tuned to the transmitter's frequency will increase as the person's body (or portions of it) get closer, building as the capacitive coupling grows stronger. If an array of receiver electrodes is placed around a person in contact with a transmitter, his or her body can be continuously tracked.

An example of the exploitation of the transmit mode is the Media Lab's Sensor Chair [Fig. 9 and Fig. 1, p. 19]. An electrode mounted under the chair's seat drives a seated body with a signal of a few volts at 50–100 kHz, well below environmental or broadcast safety limits. Four receive electrodes are mounted in a square formation in front of the player to track hand motion, and a pair on the floor below measure the position of the feet.

Halogen bulbs are mounted near the pickup electrodes and respond to the detected signals. Lights, counters, and other displays help the player during a performance by giving cue signals, clocking the time, and so on; a set of pushbuttons on the floor allow the player to fire MIDI events independent of the sensor readings [VIDEO 8].

This design can be adapted in a number of ways. For instance, the Gesture Wall, used in the *Brain Opera*, transmits the driving signal into the bodies of standing players through their shoes. Each player presents a different impedance to the transmitter, however, depending primarily on what shoes he or she is wearing. The designers therefore included in the Wall a simple servo circuit to adjust the transmitter amplitude for each player.

The Gesture Wall's receive electrodes surround a projection screen; the performer controls musical streams as the Sensor Chair player does, while at the same time interacting with the projected graphics [Fig. 10/VIDEO 9].

Shunt mode is the second electric-field–sensing mode explored at the Media Lab. Here, the body is much more strongly coupled to the room ground than to transmit and receive electrodes. As parts of the body move into and out of the region between the electrodes, the received signal level drops and rises. Increasing the body mass between the electrodes also increases the amount of signal shunted to ground and thereby decreases the level of the received signal. Decreasing the mass, naturally, does the opposite.

The nonlinear three-point coupling among transmitter, receiver, and human body can render tracking applications more computationally intensive than in simpler transmit schemes. The physical interface in shunt mode can frequently be simpler, however, because the body need

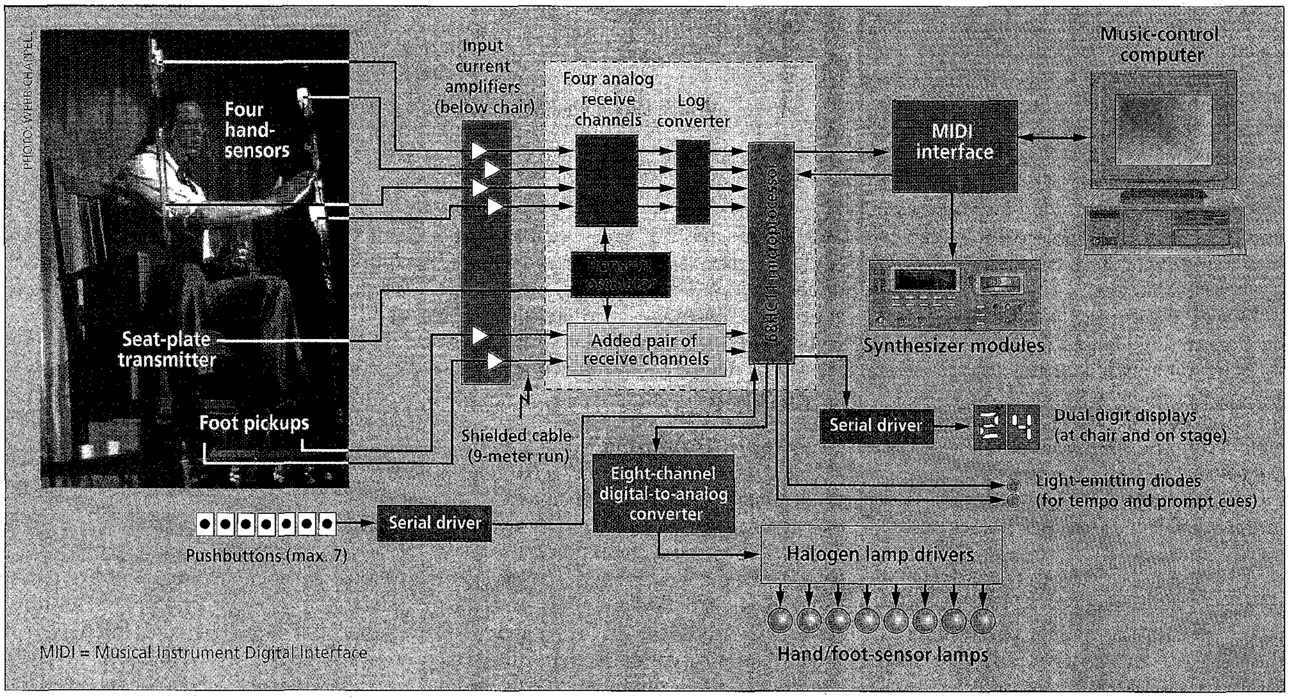


PHOTO: WEBB CHAPPEL

[9] The Sensor Chair exploits capacitance coupling of an electrically driven human body. The body is driven by a transmit signal of a few volts at 50–100 kHz, and receive electrodes are in front of the player and on the floor. Gestures of the hand and foot near the receivers send MIDI data to the computer, which controls synthesizer modules and halogen lamps. The player can also use his feet on pushbuttons to send triggers outside of the sensing field. Dual-digit displays and flashing light-emitting diodes can be used to cue the player during a performance.

not be in contact with an electrode, relying generally on its ambient coupling into the room ground.

The MIT Media Lab has created several shunt-mode musical gesture systems. One, a rectangular sensor frame, is a very simple structure well adapted for free-gesture solos and duets [VIDEO 10]. The other is a lifelike mannequin with sensors on its head, lips, limbs, joints, and one internal organ. The mannequin was custom-designed for The Artist, a rock performer formerly known as Prince. As in the other electric-field–sensing controllers described here, musical output can be mapped to a wide range of physical activity in relation to the mannequin.

Sensitive space

Noncontact detection of musical gestures using other sensing modes is being explored in several research labs and commercial products. One, the Magic Carpet, is a hybrid model designed at the Media Lab. It senses the player literally from head to toe, combining contact and noncontact modes to give the person the strong feeling of being bathed in sound. The contact part of the device senses the foot's position and dynamic pressure through a grid of piezoelectric cable running beneath a carpet. A pair of microwave motion sensors mounted beyond the carpet perimeter detects upper body gesture [VIDEO 11]. This arrangement has been employed both as a musical interface for dancers and in public installations, where passers-by explore different sonic mappings by moving through the sensitive space.

Some sonar range-finding devices that locate the user in space have been built, but most have difficulty dealing with extraneous noise, clothing-dependent reflections, and varying response speeds. Near-infrared (IR) systems avoid most of these problems, and tend to be less expensive and faster than sonars.

Although many do not work well when the players are wearing dark clothing, near-IR systems generally respond well to skin and have established a niche as hand sensors in some modern musical applications. Few of these systems measure range directly; rather, they infer it from the intensity of the reflected signal. Errors are thus introduced as a function of target reflectance.

A small number of commercial products along these lines have appeared. One is a MIDI-compatible IR hand-sensing device, the Dimension Beam, from Interactive Light, Santa Monica, Calif. Another, Synth-A-Beams, from American DJ Audio, Los Angeles, produces a MIDI event when any of eight lightbeams is interrupted.

One of the most expressive devices,

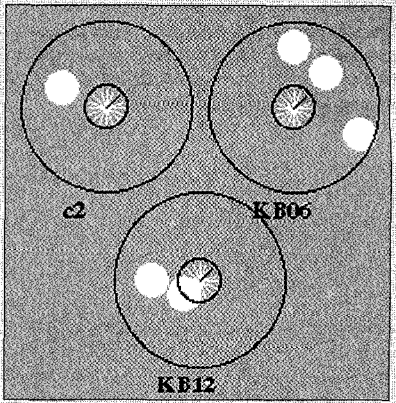
The Web resounding

Making music together: the phrase evokes visions of one of the highest forms of social expression. But even though new communities of Internet users, with avatars for stand-ins, are cropping up in what amounts to a virtual-world land boom, their musical creativity has been stymied [see *IEEE Spectrum*'s Special Report, "Sharing virtual worlds," March 1997, pp. 18–51]. The silence is now broken. On two new World Wide Web sites for Java applets, any number of players can now perform Webmusic together for a worldwide audience.

One of the sites, Spinning Disks, is currently set up for real-time jamming—just go to the Web page, and sound off with anyone else at that page. The other site, The Palette, was originally designed for Web users to add music to the Massachusetts Institute of Technology (MIT) Media Laboratory's touring extravaganza, the *Brain Opera*. The opera house, so to speak, is closed for the time being, so Palette players will not be able to join ensembles with others. However, a "practice room" is available, where the instrument can still be used for solo performance. [IEEE members have direct access to these sites through the hyperlinked version of this text at www.spectrum.ieee.org/spectrum/music/webmusic.html].

Spinning Disks was developed by Eric Metois, formerly at the Media Laboratory, and now a research engineer at ARIS Technologies Inc., also in Cambridge, Mass. His system has a graphical interface with (surprise) three large circles that look like vinyl records [see screenshot]. A smaller circle (the "label") has the same center-point as the large circle and is swept continuously by a line, much as a watch dial is swept by the second hand. At any one time, each disk is associated with—has on it—a small set of sounds, each of which is sounded with a mouse click. The sound lasts for one rotation of the moving line, but new sounds can be played at any time. Each sound per disk is on a different circumference, or "groove," of the LP, and a dot is displayed on the disk to indicate which sound has been played. (Spinning Disks is still in prototype, and the graphical representations of the grooves have not been drawn yet.)

According to its designer, Spinning Disks could be extended in many ways. Homebuilt jukeboxes are sorely needed: users should be able to add their own sonic events (timed sounds or sound sequences). In this scenario, each disk is scheduled to begin playing, at a certain tempo, and with an assigned time before sonic events are erased from it (Metois calls it "memory"). The musical experience could be supervised by a conductor, who might assign different properties to the various disks, such as speed of rotation, amount of "memory," and sonic material.



These events take place locally, at the user's computer (that is, client-side), keeping lines clear. Nor would it take much bandwidth to notify other users of changes in sonic events on disks, thus ensuring satisfactory responsiveness over the Internet.

In some circumstances, a client-side applet is inappropriate. It would simply be too limited to provide rich musical control or interact with live performances. Take the *Brain Opera*. During its performance, Internet users play The Palette, which is housed on a central server handling musical parameters. The server directs the instructions by way of a Musical Instrument Digital Interface (MIDI) stream to five computers and a battery of sound modules.

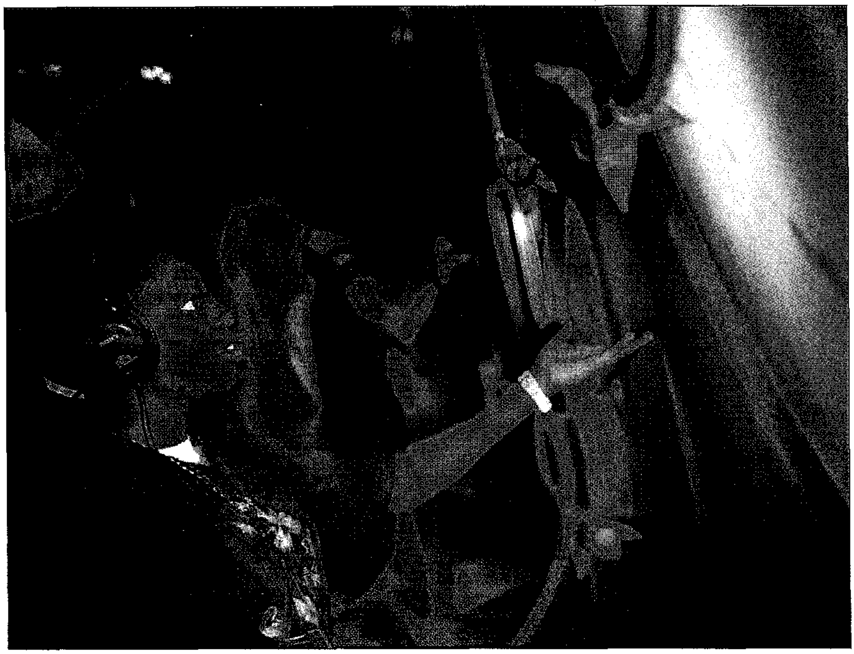
The Palette is based on the work of MIT graduate John Yu, now with Microsoft Corp., Redmond, Wash. Once on line, Paletteers can tweak 16 sonic parameters, ranging from Instrument, Key, and Pitch Direction to Rhythm Style and Rhythm Length. Perhaps the most interesting of the 16 parameters is Mutation, which controls the rate of random variation of all the other 15. The MIDI streams coming in from each Internet user were gathered, interpreted to produce the sounds, and integrated with live performance. To top it all off, the entire musical event was broadcast live over the Internet.

Stay tuned.

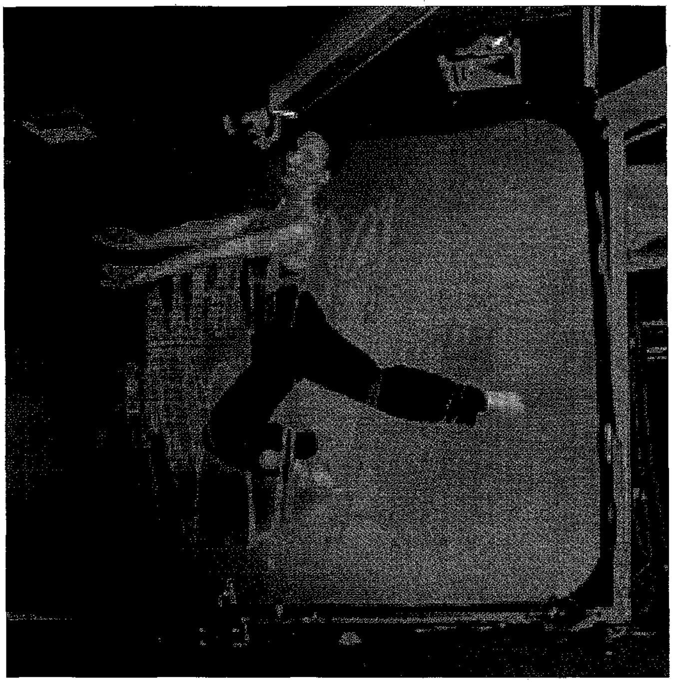
—Robert Braham

[10] The Gesture Wall [right, being explored in a performance of the *Brain Opera*] uses capacitance principles like those behind the Sensor Chair. The performer controls musical streams with hand gestures while interacting with projected graphics. A simple servo circuit in the Wall adjusts the transmitter amplitude for each player, to match the impedance of his or her shoes.

[11] DanceSpace [below, right] analyzes any body seen through its video, identifies its parts, and tracks their location. The user assigns musical controls (and graphical ones, if desired) to the features and how they are moved—for example, head height controls volume, hand positions adjust the pitch of different instruments, and feet movements fire percussive sounds.



JOHANNES KROEMER



Twin Towers, is not a commercial product. It was developed by Leonello Taraballa and Graziano Bertini at the University of Pisa. This system consists of a pair of stationary optical sensor assemblies (one used for either hand); each assembly contains an IR emitter surrounded by four IR receivers. When a hand is placed above one of these rigs, the relative balance between receiver signals varies as a function of hand inclination while their net amplitude varies with distance, indicating both range and two-axis tilt. Twin Towers resembles a Theremin but with more degrees of sonic expression arising from the extra response to the hand's angle, which is not measured in any of the other noncontact systems described here [VIDEO 12].

The Media Lab is exploring a scanning laser rangefinder, promising superior range, resolution, and robustness. This inexpensive, eye-safe phase-measuring device can resolve and track bare hands crossing a scanned plane, with updates at 25 Hz. It can sense objects up to 4 meters away and resolves much better than a centimeter in both planar coordinates.

Because the laser detection is synchronous, it is insensitive to background light; because it measures true range, errors are not introduced from varying reflectances. The lab has used this device for multimedia installations in which performers control musical and visual events by moving their hands across scanned areas above a projection screen [VIDEO 13].

Show me what you want

Computer vision techniques involve considerably more processor overhead and generally are still affected by lighting changes and clutter. But for over a decade, researchers have been designing vision systems for musical performance, while steady increases in processing capability have improved the systems' reliability, responsiveness, and ability to detect detailed features.

A straightforward example of this is the Imaginary Piano by Tarabella of the University of Pisa. A vision system using an ordinary video camera tracks the hands of a seated player. Pitch is determined by where the hands are horizontally, and a note is played when the hands pass below a certain threshold point, with pitch determined by their horizontal coordinate [VIDEO 14].

BigEye, the video analysis environment written by Tom DeMeyer and colleagues at Amsterdam's Stichting voor Electro Instrumentale Muziek (Steim), is one of the latest to be explicitly designed for live performance. Running in real time on a Macintosh computer, it tracks multiple regions of color ranges, which might, for example, correspond to pieces of a performer's clothing. Sensitive (triggering) regions of the image can be defined; different MIDI responses are specified as a function of the object's position, velocity, and so on.

More-acute machine vision software can turn the body into a musical instrument without needing specific targets or clothing. In Flavia Sparacino's DanceSpace, developed at the Media

WEBB CHAPPELL

[12] Slip into a comfortable synthesizer in the guise of this jacket stitched with conductive thread. The boat-shaped board [left] houses a single-chip synthesizer. A MIDI controller with 12 touch-sensitive "keys" lines the swatch that will be stitched on the jacket.

PETER MENZEL

[13] Composer Tod Machover models the Exos Dexterous Hand Master, a computer interface popular in virtual reality circles. The interface monitors finger position and the angle of each joint. Machover used the device to conduct his 1990 *Bug Mudra*.

Lab, a computer tracks the body when it enters the field-of-view of a digital video camera. Parts of the body—head, hands, torso, legs, and feet—are identified in real-time by the Pfinder program by Media Lab researcher Chris Wren and colleagues.

The body features can be displayed in real time on a computer monitor. For that matter, they can look like most anything, since they, too, can be remapped to any number of *graphical* events [Fig. 11/VIDEO 15]. Depending on the details of the scene, update rates on the order of 20 Hz can be achieved using a desktop workstation.

DanceSpace attaches controls to the features and how they are moved—for example, head height controls volume, hand positions adjust the pitch of different instruments, and feet movements fire percussive sounds.

Accordingly, music and accompanying graphics are generated when someone moves freely through the sensitive space. The system has been used in several dance concerts, where the dancer is unconstrained by precomposed music.

Sound swaddles

The many degrees of freedom of the human body also are captured by wearable interfaces, where sensors are affixed to the body or clothing of a performer. So far, musical applications mainly have been in interfaces for dancers (such as the MidiDancer sensor suit by Mark Coniglio of Troika Ranch, N.Y.) or performance artists (such as Laurie Anderson's drum suits). The most advanced wearable interface to have hit the commercial music world is Yamaha's Miburi system, which until recently was available in Japan.

The system is still the basis for ongoing research at Yamaha. It consists of a vest hosting an array of resistive bend sensors at the shoulder, elbows, and wrist; a pair of handgrips with two velocity-sensitive buttons on each finger; and a pair of shoe inserts with piezoelectric pickups at heel and toe [see Cover]. Current models employ a wireless datalink between a belt-mounted central controller and a nearby receiver-synthesizer.

Moreover, Yamaha invented a semaphore-like gestural language for the Miburi, in which notes are specified through a combination of arm configurations and key presses on the wrist controllers. Degrees of freedom not used in setting the pitch are routed to timbre-modifying and pitch-bending parameters. The Miburi has already figured in several virtuosic—and athletic—performances.

Stepping out

Wearable music interfaces are beginning to penetrate even fashion niches. Media Lab researchers Maggie Orth and Rehmi Post have built musical denim jackets with a touch-sensitive MIDI 12-"key" keyboard. The controller communicates over conductive thread that is embroidered directly into the fabric [Fig. 12/VIDEO 16].

Other researchers at the Lab have constructed a set of expressive footwear: dance sneakers with a suite of sensors to measure dynamic parameters expressed at a dancer's foot (differential pressure at three points and bend in the sole, two-axis tilt, three-axis shock, height off the stage, orientation, angular rate, and translational position) [VIDEO 17]. The battery-powered shoes offload their data over a wireless link.

Other wearable sensor systems developed by the virtual reality community have been eagerly adopted by musical researchers. In particular, data gloves have appeared in numerous musical performances. An example is the Exos Dexterous Hand Master, used by Tod Machover to conduct his 1990 *Bug Mudra* [Fig. 13].

Some of the most intriguing glove and hand controllers have come from Steim in Amsterdam, dating from director Michel Waisvisz' performances with his original Hands controllers in 1984. An especially complex—and thus potentially more expressive one—is Laetitia Sonami's Lady's Glove. Bend sensors measure the inclination of both finger joints for the middle three fingers, and microswitches at the end of the fingers provide tactile control. Hall sensors measure the fingers' distance from a magnet in the thumb. Pressure sensing occurs between index finger and thumb, and sonar ranging to emitters in the belt and shoe [VIDEO 18].

Perhaps the most extreme "wearable" sensors are those that respond to a person's autonomic or near-autonomic nervous systems. Some of the best-known works exploiting biofeedback were produced by David Rosenboom of Mills College, Oakland, Calif., in the '70s. The Biomuse, a commercial system from BioControl Systems in Palo Alto, Calif., works with electrical signals generated by the functioning of the muscles, eyes, heart, and brain. The Biomuse is able to produce MIDI events and has been used in musical research projects.

Although the controllability and bandwidth of some of these parameters (especially brain waves) may be debatable, new musical applications will not be far behind. Researchers at various institutes are making progress in extracting and identifying new and more precise bioelectric features.

More control needed

As it has for centuries, technology will continue changing the ways in which we create and interact with music. New developments in synthesis such as physical modeling—generating sound by mathematically simulating a mechanical or analogously complex dynamic system—demand additional degrees of control for expressive performance. In turn, this will encourage the development and acceptance of new multimodal interfaces. Further, more powerful algorithms will be employed to map real-time gestures into musical content, giving a solo performer or a networked ensemble control over highly complex synthesized sound at several levels of detail.

Because all modern electronic music controllers are digital input stations, the mechanics of electronic performance are merging with general research on human computer interfaces. In some cases, both idioms have already been combined. A case in point is composer/programmer Laurie Spiegel's Macintosh program Music Mouse, where complicated musical sound is produced and orchestrated by clicking and dragging a simple mouse [VIDEO 19]. Another is Pete Rice's recent program Stretchables, written at the Media Lab, which allows a user to perform music by drawing and dragging deformable, elastic objects, with dynamic graphical behavior strongly coupled to the sound produced [VIDEO 20].

As computer input devices improve, diffusing throughout our environment and bringing a virtual information landscape closer, we're sure to bring our musical gear along, whatever form it may take. ◆

About the author

From 1981 to 1983 Joseph A. Paradiso (M) was at the Laboratory for High-Energy Physics at the ETH in Zurich, and from 1984–1994 he was at the Draper Laboratory, Cambridge, Mass., where he conducted research in spacecraft control systems, image processing, sonar sensors, and high-energy physics detectors for the Superconducting Supercollider. Since 1994, he has been a research scientist in the Physics and Media Group at the Massachusetts Institute of Technology's Media Laboratory. He is currently the technology director of the Things That Think consortium, a collaboration between Media Lab researchers and industrial partners exploring computing, communication, and sensing technologies embedded in commonplace environments.

Spectrum editor: Robert Braham

To probe further

A greatly expanded version of this article will be posted next month on the author's home page at the Media Laboratory of the Massachusetts Institute of Technology, http://physics.www-.media.mit.edu/~joep/

A good history of electronic music is *The Art of Electronic Music, by Eds.* Tom Darter and Greg Arnbrysterm (GPI Publications, N.Y., 1984).

Joel Chadabe's *Electronic Sound: The Past and Promise of Electronic Music* (Prentice-Hall, N.J., 1997) is an excellent summary of the current state of instruments, emphasizing innovators at research centers.

A hefty and up-to-date tour of computer music, including a chapter on interfaces, is *The Computer Music Tutorial*, by Curtis Roads (MIT Press, Cambridge, Mass., 1996). On the Musical Instrument Digital Interface (MIDI) protocol, an excellent introduction is Paul D. Lehrman and Tim Tully's *MIDI for the Professional*, 2nd ed. (Amsco Publications, New York, 1995).

Technical summaries of the *Brain Opera* and other Media Lab research in this area can be reached starting at physics.www.-media.mit.edu/~joep/TTT.BO/index.

Electric field sensing, both its history and how it is used in Media Lab instruments, is detailed in "Musical Applications of Electric Field Sensing," by the author and Neil Gershenfeld, *Computer Music Journal*, Vol. 21, no. 2, Summer 1997.

Details on the Media Lab's optical tracking interfaces, including the Digital Baton, laser rangefinder, and DanceSpace, are in "Optical Tracking for Music and Dance Performance," by the author and Flavia Sparacino, published in *Optical 3-D Measurement Techniques IV*, Eds. A. Gruen and H. Kahmen (Herbert Wichmann Verlag, Heidelberg, 1997), pp. 11–18.

A bestiary of historical electronic instruments, with some sound files, is on the World Wide Web site at www.obsolete.com/ 120_years/; www.ief.u-psud.fr/~thierry/history/history.html is another good history site. And definitely worth renting is the video of Steve Martin's 1993 documentary *The Electronic Odyssey of Leon Theremin*, Orion Pictures, 1993.

For links to Musical Instrument Digital Interface (MIDI) controller sites, the best place to start is www.synthzone.com/ctrlr.htm. A technically savvy list-of-lists is capella.dur.ac.uk/doug/music.html, which is maintained by Douglas Nunn of the Durham Music Technology Group at the University of Durham.

More information about research groups mentioned in this article is available at their Web sites:

- The Center for Computer Research in Music and Acoustics, www-ccrma.stanford.edu/.
- The Center for New Music and Audio Technologies, www.cnmat.berkeley.edu.
- Institut de Recherche et Coordination Acoustique/Musique, www.ircam.fr/index-e.html.
- The Stichting voor Electro Instrumentale Muziek, www.xs4all.nl/~steim/.
- The music technology center at the University of Pisa, spcons.cnuce.cnr.it/music/cmd.html.

The IEEE focus group on electronic music can be reached at www.computer.org/tab/cgm/tc_cgm.htm.