From*Proceedings of the 1995 Connecticut College Symposium on Arts and Technology*

Strategies for Interaction: Computer Music, Performance, and Multimedia

Todd Winkler
Brown University

ABSTRACT

Interactive music is defined as a music composition or improvisation that interprets human actions to affect some aspect of music generated or modified by computers. Typically, this involves a performer playing an instrument while a computer creates music that is, in some way, shaped by the performance. Interactive music techniques developed over the past few years can apply to multimedia interactivity in the arts, in CD-ROM titles, installations and performance works. In screen-based works, the user (the “performer”) may control aspects of music selection and compositional processes using the computer keyboard and mouse. Continuous control over musical parameters adds a level of variation and interaction not usually seen in commercially available titles. Techniques for analyzing performance action may also be applied to trigger or modify visual elements, such as text or video, in a concert presentation using computer projectors and large monitor systems.

Music has always been an interactive art full of collaborators who actively participate together: the conductor and the orchestra, the lead singer and rock band, or a jazz ensemble trading solos. Performers listen to each other while they play, reacting all of the time to what they hear. Musical performances usually stay within a consistent context that creates a feeling of understanding among the musicians and the audience without being entirely predictable.

Interactive music is a natural extension of a long history of collaborations. Interactive music uses software to interpret human action to effect some aspect of music generated or modified by computers. Performer’s actions directly or indirectly alter parameters that control computer music processes. Performance information may be used immediately to alter compositional algorithms, or may be viewed over longer spans of time, such as analyzing phrases or sections of a work. Any analysis data may be stored for future use. Typically, data input involves a performer playing an instrument that sends MIDI data, while a computer creates music that is, in some way, shaped by the performance. The computer keyboard and mouse, or any other hardware device that translates motion into computer data, may also be used as input devices to shape musical processes. Thus, while most of the research and concepts discussed here relate directly to musical performance, the idea of a performer may be expanded to include anyone capable of influencing musical decisions in computer-based artwork. Many of the thorny technical issues surrounding interactive art have been ironed out by music researchers over the past ten years. These techniques give performers the feeling of participating, along with the computer, in the creation of a musical work. The feeling of interactivity depends on the amount of freedom the performer has to produce and perceive significant results [Laurel, 1992].

On the low end of the interactivity scale lies predetermined sequences or sound files that are triggered by a performer. This is the typical technique found in CD-ROM titles and computer games. While robust branching structures and mixing capabilities allow the performer (user) to experience musical selection, the samples or MIDI files, when triggered, produce the same results. (This is about as interactive as an audio CD player). A predetermined score could be made slightly interactive by allowing a performer to control only one parameter of the music, such as the tempo or dynamics. Highly interactive pieces are less predictable since the performer controls many more significant musical parameters, and in turn, reacts to the computer’s response. The composition can change dramatically according to a performer's actions and interpretation. Melodic construction, harmony, tempo, register, dynamics, and timbre can all be influenced in real-time by performance information. In the most robust examples, a musician is free to play any kind of music, and the software has enough “intelligence” to respond in a way that makes sense and naturally elicits the performer’s continuation. Like a good conversation, interactive music succeeds when it encourages spontaneity while residing within the boundaries of a dynamic musical context that is whole and engaging.

Interactive software simulates intelligent behavior by modeling human hearing, understanding, and response [Rowe, 1991]. The response must be believable in the sense that it seems appropriate for the action taken, and appropriate for the style of music. This process is somewhat analogous to the distinct activities that take place simultaneously during a jazz improvisation or other musical dialogue: listening, interpreting, composing, and performing.

Computer Listening

At the heart of interactive composition lies software capable of "understanding" musical performance. Musical understanding, both in humans and computers, occurs when an event or series of events can be matched in memory with a similar event or with a concept of the event [Laske, 1980]. This could be the recognition of simple data such as "Middle C," the concept of a spatial relationship, such as a "major chord," or a process, such as "slow down." Musical memories are referenced to provide insight into the listening situation of the moment.

Software can be equipped with the memory (knowledge) of basic musical elements, such as notes, dynamics, and duration, or larger musical events, such as phrases, riffs, melodies, and rhythmic patterns. "Listening" algorithms analyze and store a wide variety of common musical events, from relatively simple data input. A typical MIDI system, for instance, usually transmits pitch and loudness (velocity) information. With just these two parameters as input, a computer can measure how they change over time revealing an abundance of useful musical data, such as chord structure, key, rhythm, phrase length, and increases and decreases in tempo and dynamics. The computer keyboard and mouse represents a similarly limited input mechanism that can be interpreted to produce complex musical results, with listener algorithms designed to detect patterns and timing information.

Computer Composition

Flexible compositional algorithms may generate an infinite variety of music, with selected parameters influenced by performers. For hundreds of years, composers have used such processes to generate large musical structures from simple musical material. In interactive computer music, these generative processes, in the form of algorithms, use basic musical material, either previously stored or derived from the performance. Changeable parameters allow the performers to continuously shape and influence musical output. The addition of constrained random variables sets up improvisational situations where the music will always be different and unpredictable, yet it will always reside within a prescribed range that defines a coherent musical style.

Current research shows promise that future versions of personal computers will be fast enough to have direct sound synthesis and digital signal processing capabilities, thereby replacing the need for external MIDI devices [Lindeman, 1991]. This will enable the performer to shape timbre as well as other musical parameters. With sample playback, signal processing, and variable synthesis algorithms available for sound generation from the computer, interactive music techniques will become more viable for a large number of artists since the music will not depend on specific hardware devices and will therefore travel easily from one computer to another.

Multimedia Extensions

Many composers, already experienced with creating artistic work using the computer, will find that expanding music into the multimedia realm is a natural extension of their ongoing work. Because music is fundamentally a time-based art, composers are well trained to deal with the aesthetic issues surrounding timing structures. The inclusion of QuickTime extensions into many computer music programs, most notably Max [Puckette and Zicarelli, 1990], creates opportunities for using flexible musical structures to interact with video, still images, animation and text. Max is ideally suited for playing video and still images in a performance situation, or in any application where timing functions are critical. Because Max has the extensive and reliable real-time capabilities required for musical production, it is rapidly becoming an alternative language for visual artists and musicians who wish to incorporate multimedia elements in their work, offering a fascinating environment to explore the combination of sound and images that are immediately responsive to human gestures.

For live performance, music software has some advantages over other programs since it is already optimized for immediate response and feedback. Max has been used for extensive real-time video playback systems developed by AVX, and used in multi-screen video performances by Electric Broadcasting Network (EBN). Other artists are using Max for installations to control playback of images from videodisc players.

For new media titles viewed on the computer, interactive music techniques add realism and user participation in creating effective soundtracks by using compositional algorithms that change continuously in response to a story or in response to the user's actions. Hyperfiction and CD-ROM titles based on the Hypercard paradigm, usually present static scenes that offer user interaction in the form of mouse clicking to select one of several options. Repetition seems inevitable in these situations as the user stumbles back to previous scenes. The success of *Myst* [Miller, 1994] demonstrates how two devices can make this repetition bearable and even interesting. First, the images and sounds can be complex and rich in detail, so that on a second or third look new aspects are discovered. Second, upon returning to scenes that appeared to be static, images and sounds can be transformed to reflect changes in the story. The next step, continuous control of parameters, has already taken place in the visual domain, with continuous 360 degree viewing available in a recent QuickTime addition, QuickTime VR.

Giving audio and soundtracks such flexibility would add a great deal of interest. For example, timing of decision making, mouse movements, or keyboard strokes, could be linked to control the tempo of the music; or, clicking and dragging on object downward might cause the music to go lower, or the timbre to get darker. Constrained random values could be used to create endless variations, so that the music for one scene would always be changing but have the stylistic consistency required of the dramatic structure. Parameter changes could also be controlled by the time spent within a single scene, so that the music would intensify as the viewing continued. Clearly, music and sound responsive to user input will increase the feeling of interaction and participation.

New Performance Devices

Numerous controllers may be used as input devices to translate gestural information into computer code sending either discreet data (e.g. MIDI keyboard or computer keyboard) or continuous data (e.g. MIDI modulation wheel, or computer mouse). Audio signals from a voice or instrument may be translated into MIDI data using a pitch-to-MIDI convertor. The increase in sophistication of new MIDI controllers that allow physical gestures and movements to be translated into MIDI data holds great promise for interactive composition and for creating responsive musical scores for multimedia pieces and virtual reality systems. Freed from the physical limitations of acoustical instruments, many new digital controllers have been developed as idealized motion transducers. The new controllers, such as Mark Coniglio's *MIDI Dancer*, Donald Buchla's *Lightning*, and Max Mathew's *Radio Drum* (to name a few) have several things in common: they all specify a limited range of motion (left-right, high-low, open-closed), divide that range into a limited number of discrete steps, and send out data more or less continuously along their range. Numbers, often represented as MIDI continuous controller values between 0 to 127, determine location over time within this pre-defined range. These series of numbers can be used as variables to influence compositional algorithms that produce musically satisfying results. New controllers send out numbers that represent physical location in a room, arm motions, hand position, eye movement, and even neurological activity in the muscles [Knapp and Lusted]. Although most of the devices described here were invented for musical performances, QuickTime extensions make them equally adept at controlling visual elements. Similarly, physical interfaces for the computer, not intended specifically for music, may also be used to create soundtracks or used as stand alone instruments. Any device that translates physical movement into numbers is fair game for an interactive musical performance.

New physical controllers are hardware devices directly played with parts of the body. Some of these instruments measure movement and position of the arms, legs, and especially hands and fingers directly converting this into information to make music. Michel Waisvitcz's *The Hands* [Waisvitcz, 1985] is one of the first such devices. It looks like a pair of oversized gloves with switches and fits on both hands, translating finger position, hand proximity and wrist motion into MIDI data. Tod Machover's Dexterous Hand Master, and other devices are based on a similar concept [Machover, 1989]. Even game and toy makers, notably the Mattel Powerglove, have come out with inexpensive glove controllers which have been used for interactive music performances. Mark Coniglio's MIDI Dancer, is a wireless device consisting of a thin elongated hinge that translates the angle of arm and leg joints into MIDI data. The device uses Coniglio's *Interactor* software to create music based on dancers' arm and leg position, motion, and speed. In his latest piece, *In Plane*, a solo dancer's role is expanded to control video playback, a lighting board, and the position of a video camera on a 24 foot robotic platform [Coniglio, 1995].

Motion detectors respond to the location and movement of a performer in space, without the performer touching any hardware devices. The Theremin, an instrument built in 1919 by Leon Theremin, is the earliest such example. It is played by waving the hands near two antennae, with the proximity of one hand controlling frequency, the other amplitude. Several performers have used off-the-shelf motion detectors as well as custom-designed hardware to map body location within a space into MIDI data. Donald Buchla's Lightning uses infrared waves to map location from a wand onto a user-definable grid that outputs MIDI values [Rich, 1991]. David Rokeby's *Very Nervous System* [Cooper, 1994] tracks a person's movements in a large space using a video camera to send images to a computer which analyzes consecutive frames to detect motion. Using Max software to create music, the system allows dancers and other participants to influence compositional processes based on their location and movement in the room. The same system has been used for video artist Paul Garrin's twelve monitor interactive video installation, entitled *White Devil*.

Conclusion

Interactive music suggests a new paradigm for computer music composition; one where the very fabric of a composition is informed by the subtleties of human gesture. These techniques offer non-musicians the direct experience of creatively shaping music through flexible compositional algorithms. At the same time, similar techniques can be used to capture musical nuances from expert musicians whose skills may impart a human expressive element to machine made music. The extension of these techniques into the visual domain holds great promise for future developments in multimedia performances and CD-ROM titles.

Referernces

Detailed explanations of the musical concepts presented here will be found in the author's forthcoming book entitled, *Composing Interactive Music: Ideas and Techniques Using Max*. Madison, WI: R-Editions, Inc, planned for 1996.

Coniglio, Mark. *In Plane*. Personal conversation. 1995.

Cooper, Douglas. "Very Nervous System." *Wire Magazine*. 3, 3 (1995): 134-170.

Knapp, Ben, and Hugh Lusted. *Biomuse System*. Personal conversation.

Laurel, Brenda. *Computers as Theater*. Reading, Mass: Addison-Wesley Publishing Company, 1993.

Laske, Otto. "Toward an Explicit Cognitive Theory of Musical Listening." *Computer Music Journal* 4, 2 (1980): 73.

Lindemann, Eric; Dechelle, F.; Smith, B.; and Starkier, M.. "The Architecture of the IRCAM Musical Workstation. *Computer Music Journal.* 15,3 (1991): 41-49.

Machover, Todd, and Joseph Chung. "Hyperinstruments: Musically Intelligent/Interactive Performance and Creativity Systems." In *Proceedings of the 1989 International Computer Music Conference*, ed., T. Wells and D. Butler. San Francisco: Computer Music Association, 1989.

Miller, Rand and Robyn Miller. *Myst*. Spokane, WA: Cyan Inc., 1994.

Puckette, Miller and David Zicarelli. *MAX - An Interactive Graphic Programming Environment*. Menlo Park: Opcode Systems Inc, 1990.

Rich, Robert. "Buchla Lightening MIDI Controller." Electronic Musician 7, 10 (1991): 102-108.

Rowe, Robert. *Interactive Music Systems.* Cambridge, Mass.: MIT Press, 1992.

Waiswicz, Michel. "THE HANDS: A Set of Remote MIDI Controllers." In *Proceedings of the 1985 International Computer Music Conference*, ed., B. Truax. San Francisco: Computer Music Association, 1985.