## Gestural Control of Computer-Based Musical Instruments

No volume that is devoted to a discussion of new interfaces for musical expression would be complete without a discussion of Gestural control. Loosley defined, gestural control addresses the question of how to map the capabilities of the human sensorimotor system to the parameter space of the instrument being played, in order to provide the performer with maximum control of music's four fundamental elements - time, pitch, timbre and amplitude. For the performer, manipulation of these dimensions is embodied in physical actions such as striking a piano key or bowing a string. In real instruments these ``instrumental gestures'' are determined by the sound-producing mechanisms of the instrument. The sound produced, in turn, carries the characteristics of the movement that gave rise to it[1].

A by-product of sound synthesis techniques such as physical modelling [2] is the ability to decouple the synthesis of an instrument's sound from the physics of the instrument's sound-producing mechanism. Thus the affordances of a synthetic music controller can be very different from those of the instrument being controlled. In other words, there is no longer a direct mapping between the instrument and the instrumental gesture. A percussive controller such as a piano keyboard, for example, might be used to control a physical model of a bowed string. The question then arises: What are the implications for the design of computer-based musical instruments and their supporting hardware and software protocols of decoupling the instrumental gesture from the sound synthesis model? While it is impossible to address this question in full within the scope of these pages, it is hoped that the issues raised here can motivate a debate about the representation of gesture in a control protocol for computer-based musical instruments.

## MIDI and Gestural Control

It is almost an accident of history that early music controllers were predominantly based on the piano keyboard. Keyboards were amongst the first electronic music controllers to be developed, possibly because they could be implemented using simple and reliable electronic components such as switches. An unfortunate by-product of the prelifferation of keyboard controllers for MIDI was that the MIDI protocol was early biased toward the affordances of keyboard controllers. The most obvious consequence of this bias is a lack of support in MIDI for a control hierarchy that allows note-level events to be subordinated to phrase-level control parameters. For keyboard instruments, more than for any other instruments (except perhaps percussion instruments), there exists a one-to-one mapping between a note and the movement that produced it. Thus the development of the MIDI protocol was predicated on the assumption that each note is an isolated event with controls for its pitch, duration, timbre and amplitude, that cannot interact. For most musical instruments, this one-to-one mapping is the exception rather than the rule. For bowed instruments, for example, many notes of a single slurred phrase will be executed with one bow stroke and hence a single arm movement. Not only does this single arm movement cause notes of a phrase to be linked by a common bow stroke, but its trajectory also embodies expressive nuance, shaping the dynamic and timbral arc of the phrase. Parameters for dynamics and tone color are thereby correlated and co-vary in response to the trajectory of the movement of the arm.

To capture such nuance in performance and convey its subtleties to a synthesis model requires a control architecture where the gesture, not the note, is the primary unit of musical time. Several researchers (most notably [1, 3, 4]) have proposed that music controllers and the protocol that supports their communication with synthesis algorithms should be founded on an asynchronous hierarchical structure with the performance gesture, not the score-based note list, as its unit of currency. As [4] points out, players of non-keyboard instruments have been reluctant to embrace the digital revolution in music. Woodwinds, bowed strings, and brass instruments all place the player in direct physical contact with the vibrating element --- reeds, strings, or columns of air --- providing the player with fine control of a single note or a small group of notes. Most commercially available control devices provide limited control of multiple notes and are inappropriate for most melodic instruments. Faced with trading fine control of a real instrument for the infinite timbral possibilities but coarse control of today's synthesis models, most players opt to remain in the real world. Even in those cases where real instruments are adapted to transmit MIDI messages, so-called hybrid controllers, the limitations of MIDI still presents significant bottleneck. In terms of communicating musical nuance, MIDI breaks down at many levels. Purely in terms of hardware, it cannot support isochronous high-speed two-way communication between controller and instrument. But more importantly, even if hardware constraints were removed, the command structure of MIDI pays little attention to the hierarchical nature of musical performance. In music performance, expressive musical ideas, and their attendant sequences of movements, are organized by the human motor system into hierarchies. Low-level events, such as the execution of individual notes, are presumed to be encapsulated in motor programs, which can be triggered in sequences, or patterns that are also learned. But these higher-level patterns can be shaped and reshaped in real time by meta-level movement control [3]. Thus rubato, ritardando, and accelerando gestures can be superimposed at will on music that has already been learned. For a protocol to successfully support gestural control, therefore, there must exist an architecture to support this hierarchy, taking into account the ``connections'' between various control layers. Because the command structure of MIDI ignores this hierarchy of control in music, it is not capable of supporting the translation of performance gestures into musical nuance

**Toward a Protocol for Expressive Musical Control**

What, then, are the salient issues in designing a protocol that can support new interfaces for musical expression? The goal of any protocol for communicating between a music controller and a synthesis module is to support the translation of musical nuance, expressed as movement, into parameters for controlling sound. Thus the control protocol becomes an interpreter, mediating between the performer's intent and the synthesis module's realization. The question of how to design such a protocol has vexed computer music research for almost 30 years, but has become more prevalent in an age where real-time gestural control is the de facto standard for performance systems. It is timely, therefore, that we recognise the importance of a discussion of new interfaces for musical expression in their own right and, I suggest, equally timely that we think of how to support these new instruments with a protocol that captures the richness of performance nuance and makes it possible to map this to synthesis models. In motivating the discussion, I suggest that a new protocol to support the interconnection of synthesis modules and expressive controllers should have a hierarchical structure that can simultaneously represent both low-level movement trajectories and higher-level musical articulation trajectories Such a gesture-driven protocol would make it possible to correlate parameters that are coupled by the same physical movements and would also give rise to a language for gestural control. Then, finally, there would exist a context within which to explore a whole realm of issues around the mapping of physical gesture to musical control.

Sile O'Modhrain
Chairperson of the NIME-02 Local Committee

[1] C. Cadoz, “Instrumental gesture and musical composition,” in Proceedings of International Computer Music Conference, 1988.
[2] J. Smith, “Principles of Digital Wave\-guide Models of Musical Instruments,” in *Applications of Digital Signal Processing to Audio and Acoustics*, M. K. a. K. Brandenburg, Ed. Boston/Dordrecht/London: Kluwer Academic Publishers, 1998, pp. 417-466.
[3] L. H. Shaffer, “Attention: Selection, awareness, and control: A tribute to Donald Broadbent,” , A. B. a. L. Wisekranz, Ed. Oxford: Clarindon Press, 1993, pp. 135-151.
[4] K. McMillen, “ZIPI: Origins and Motivations,” *Computer Music Journal*, vol. 18, pp. 47-51, 1994.