Improvisation with Highly Interactive Real-Time Performance Systems

David Wessel
Center for New Music and Audio Technologies (CNMAT)
Department of Music
University of California at Berkeley
1750 Arch Street
Berkeley, CA 94709
wessel@cnmat.berkeley.edu

Abstract

Improvisation can make special demands on a performance computer music system, especially when more than one musician is involved. Here special emphasis will be given to the use of computer based technology to enhance the musical dialogue between two musicians. The software environment, written in MAX, is organized into objects that function as listening assistants, composing assistants, and performance assistants. Listening assistants extract musically meaningful information from the on-going performance; they include DSP based objects for pitch extraction and parsing of the acoustic data stream, as well as, tempo extractors and objects for musical analysis. Composing assistants help the improvisers manage the construction and set-up of his performance materials in an on-the-fly manner. Performance assistants help with gesture mapping and phrasing and articulation.

Design issues for systems that require the collective action of performers are also described. Demonstrations taken from actual performances will be provided.

Introduction

Improvisation, be it open ended or constrained, requires that its musical materials be malleable and capable of immediate adjustment. This requirement places special and severe demands on the design of performance oriented computer music systems. Attention must be given to mental or machine representation of the musical material, expressive real-time control, resource allocation, and scheduling. When more than one musician is involved additional attention must be given to the nature of the interaction. In this brief presentation I would like to describe some features of software that I have developed for duo improvisations and along the way touch on the just mentioned design issues.

The initial and primary design goal was to enhance interaction between two musicians in such a way that a new form of musical discourse might emerge. The computer was viewed as an intimacy amplifier in the sense that it could pay detailed attention to the performers and provide a flexible representation of the materials at hand and their musical context. Another design goal was to privilege one of the unique features of a computer mediated performance system, namely that the performers gestures can be used to control the behavior of entire phrases or processes rather than just single note events as is the case with most traditional acoustic instruments.
Gathering Phrases and Performing their Transformations

The first systems I will describe were developed around a series of performances that began in the fall of 1986 in collaboration with saxophonist Roncone Mitchell and later with Ushio Torikai on J-Muize. The basic idea was to have the musician playing the computer part use materials gathered on-the-fly from the performance of the other who was improvising freely.

A pitch extractor connected to the acoustic instrument provided a monophonic stream of data to performance software written in MAX (Puckette & Zicarelli 1990). A standard MIDI keyboard managed the recording of phrases, their transformation, and performance. The keyboard was configured so that the lowest octave functioned as a phrase recorder. Recording was initiated when a key in this octave was depressed and stopped when it was released. The stored fragment was associated with that key and the corresponding keys in the other octaves of the keyboard were used to playback transformations of the recorded phrase. In this paper only a few of the features of this performance software will be discussed and demonstrated.

Listening Assistants

When a performer decides to record a fragment from the ongoing flow of the other musician it is already too late. The phase may have already begun and the performer must somehow reach back in time to catch the start. Some form of short term memory and a phrase parsing mechanism are essential here. Further information about the phrase and its context is required for its intelligent transformation. Towards these ends a number of software elements were developed that assist in the extraction of musically meaningful information.

At the core of these listening assistants is a model of memory influenced by memory work in cognitive psychology. The shorter-term memory mechanisms operate at a level more directly related to input signals while the longer-term memories store more abstracted forms of the input material. At the lowest level the short-term memory consists of a buffer to hold the last second or so of input. A phrase boundary detection mechanism operates on this short-term store.

The design of the phrase boundary detector was based on the elementary grouping mechanisms of Lerdahl and Jackendoff (1984). Here the essential idea is gap detection. A time gap between successive notes indicates a phrase boundary. Similarity a phrase boundary is marked by an accent (an amplitude gap), a register shift (a pitch gap), and a significant timbre change (a gap in timbre space).

While experimentation has been carried out on the full implementation of these Lerdahl and Jackendoff gap-based grouping rules, only the the time-gap rule has been used effectively in actual performances. The problem and interest in using the full set of gap rules is that the different rules often give conflicting results. Experiments with these ambiguous situations have taken two different directions, one uses weighting schemes on the various sources of evidence to arrive at a single result, the other carries the ambiguity to higher levels in the system maintaining multiple representations of the grouping structure. This later approach appears to be more interesting and useful in that these multiple representations can shed light on the metrical and other aspects of the musical organization.

The demand for flexible representations of musical material with the possibility of immediate alteration in performance require that transformation and playback of the recorded phrases be available within milliseconds of the beginning of the recording process. This requirement for nearly simultaneous playback and recording is not difficult to implement in the MAX environment. On the other hand, postponements take time. Take for example the fact that one cannot make a proper inference about the metrical structure until at least one metrical unit has been played in full. Indeed, managing musical abstraction mechanisms in hard real-time systems is one of the most challenging aspects of this work on systems for improvisation.
Transformations Informed by Local Context

One of the transformations of the performance-gathered phrases that has been of interest is the generation of chords from these materials. A first approximation involves projecting the notes of the melodic line into a harmony, but this simple verticalization of a linear structure very often produces chords that are too dense and inappropriate to the local pitch context. We have experimented with a listening assistant that keeps a running tabulation of the local pitch hierarchy. This tonal field estimation scheme is based on a method developed by Krumhansl (1989). The program maintains a histogram of the relative accumulated durations of the pitches weighted backwards in time so that notes in the immediate past have more influence than those in the more distant past.

These Krumhansl-style pitch histograms were used in the chord generation process in a variety of ways. Their simple use involved using the histogram for the recent past to help edit the chords generated from the captured line. Tones with low frequency of occurrence in the estimated tonal hierarchy were eliminated from the chord. A more elaborate use involved using the local context as a point of departure for a modulation to a different key-like tonal hierarchy. Here the MAX objects for composing assistance were put to use.

Performance Assistants

One example of a performance assistant object will be given and is taken from a collaboration with Jim Hi Kim in early 1991. Here the idea was to launch a sequence with an expressive gesture made by moving a finger on a pressure and location sensitive control strip like those available on Buchla's Thundie. The phrase began when the gesture was initiated but its shape was expanded in time and applied over a duration only times the duration of the gesture itself. The gesture function was stored in MAX's table objects and the playback mechanism scanned this gesture function at a much slower rate. Various mappings of the pressure and location parameters were explored. One of the most satisfactory mappings was to control tempo by the rate of location change and the dynamic contour by pressure. This gesture-based phrase launching mechanism might be thought of as something analogous to the throwing of a Frisbee where the initial launch influences the tang term course of the flight.

Collective Control

Interactive performance systems offer the intriguing possibility of two or more musicians jointly controlling a single musical structure. Though improvisation did not play a significant role, Stockhausen's Microphonie I provides a compelling example of such collective control. Here one of the performers excises a large tam-tam with small implements. Two other performers probe and amplify this low dynamic level material with microphones whose output is further subjected to electronic transformation. The resulting sound world is not just the additive combination of the sounds generated by the individual players as in a traditional ensemble. Rather, it is the result of a cooperative interaction among the performers.

Experimentation with performance networks that require joint action to produce a sonic result is to be encouraged. Furthermore, I believe that such musically oriented groupware is best explored in an improvisational context. This is in no way rules out the role of compositional design. The challenge is to design performance paradigms that privilege genuine intimacy between the performers. Development will not come easily. Music can be a very solitary activity. Most practicing is done alone and even in cultures where ensemble play is emphasized, the individual musicians each produce their own sound. In performance paradigms where collective action is required, the elementary development of the art requires the presence of all the musicians. Furthermore, the design of the software environment requires a merger of communications and computing where the goal is to produce a performance paradigm that facilitates responsiveness not only between each musician and the instrument but also between the musicians as well.
Sonic and Extra-sonic Musical Communication

In developing such collective performance systems, forms of contact other than those based on sound should be explored. The performance network should allow for the silent communication of intentions and desires. Such multimedia forms of communication are not new to music. Visual signs abound and even in the most freely improvised music the open-eyed player gains an advantage of better reception. At this point I would like to admit to a bias against the use of screens in performance. It would seem the watching the screen often precludes the watching of the other performers. A number of the newer display technologies might be worth investigating. It would seem desirable to have a display that allows one to see what is going on around them and see simultaneously presented computer generated information. Tactile transducers might be effective as well.

Conclusions

Network performance systems blend a number of computational and communications issues.

- Large quantities of diverse material and processes must be rapidly available.
- Listening and abstraction mechanisms must be integrated into hard real-time systems.
- Composing and planning of performance actions must be facilitated.
- Expressive gestures must be mapped over entire phrases rather than just single notes.
- Musical performance groupware should be explored.
- Multimedia communications networks are essential to a network of performers who want to communicate their intentions and desires as well as the sound material.

References

