

# Visual Feedback in Performer-Machine Interaction for Musical Improvisation

Alexandre R.J. François  
Viterbi School of Engineering  
Univ. of Southern California  
Los Angeles, CA, USA  
afrancoi@usc.edu

Elaine Chew  
Viterbi School of Engineering  
Univ. of Southern California  
Los Angeles, CA, USA  
echew@usc.edu

Dennis Thurmond  
Thornton School of Music  
Univ. of Southern California  
Los Angeles, CA, USA  
thurmond@usc.edu

## ABSTRACT

This paper describes the design of *Mimi*, a multi-modal interactive musical improvisation system that explores the potential and powerful impact of visual feedback in performer-machine interaction. *Mimi* is a performer-centric tool designed for use in performance and teaching. Its key and novel component is its visual interface, designed to provide the performer with instantaneous and continuous information on the state of the system. For human improvisation, in which context and planning are paramount, the relevant state of the system extends to the near future and recent past. *Mimi*'s visual interface allows for a peculiar blend of raw reflex typically associated with improvisation, and preparation and timing more closely affiliated with score-based reading. *Mimi* is not only an effective improvisation partner, it has also proven itself to be an invaluable platform through which to interrogate the mental models necessary for successful improvisation.

## Keywords

Performer-machine interaction, visualization design, machine improvisation

## 1. INTRODUCTION

This paper describes the design of *Mimi*, shown in Figure 1, a multi-modal interactive musical improvisation system that explores the role of visual feedback in performer-machine interaction. The project results from the playful interactions of a computer scientist/amateur pianist (François), an engineer/concert pianist (Chew), and a classical/improvising keyboard performer/pedagogue (Thurmond).

*Mimi* is, first and foremost, a performer-centric tool designed for use in performance and teaching. Interactions with the system occur in two distinct phases: preparation and performance. Interaction in the two stages occur over identical modalities, but under different external constraints. In the preparation phase, a musician lays down the mate-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

NIME07, June 7-9, 2007, New York, NY

Copyright 2007 Copyright remains with the authors.



Figure 1: Aural and visual performer-machine interaction: Dennis Thurmond and *Mimi*.

rial that will be used by the factor oracle-based system for generating improvisations. This material from the human performer/improviser may result from a spontaneous creation, very similar to aspects of the improvisation process during human-machine interaction, or it may result from a carefully planned process, more akin to composition, which has sometimes been likened to slow improvisation.

From a technical point of view, *Mimi*'s system requirements span traditionally disjoint sets that are difficult to reconcile. First, use in live performance demands real-time, interactive and synchronous synthesis and display of multi-modal output. Second, the symbolic data structure (factor oracle) and algorithms that constitute the core of the improvisation engine operate outside of time, and must be seamlessly integrated with the real-time performance aspects of the system, without compromising their power and flexibility, or the quality of the real-time experience. The meeting of these requirements is facilitated by the use of François' Software Architecture for Immersipresence (SAI) framework [6, 7].

This paper presents *Mimi*'s visual interface design, and findings from initial case studies on the use and effectiveness of visual feedback in improvisation planning and design. The remainder of this paper is organized as follows. Section 2 places *Mimi*'s approach in context with past and ongoing work in interactive machine improvisation. Section 3 describes *Mimi*'s visual interface design. Section 4 offers a discussion of the system, and of the principles underlying the design of its visual interface. Finally, Section 5 offers concluding remarks, and outlines research and development directions for ongoing and future work.

## 2. MACHINE IMPROVISATION

This section relates *Mimi* to representative past and ongoing work in interactive machine improvisation. The designs of these systems rely almost exclusively on real-time auditory feedback for the performer(s) to assess the state of the system, and therefore its possible evolution in the near future.

George Lewis has been creating, and performing with, algorithmic improvisation systems since the 1980s. One example is *Voyager*, whom Lewis describes as a “nonhierarchical, interactive musical environment that privileges improvisation” [8]. In Lewis’ approach, the preparation process consists of the creating of a computer program. During the performance, the program automatically generates responses to the musician’s playing, as well as new material. The performer(s) *listens* and reacts to the system’s musical output, closing the exclusively aural feedback loop. The musician has no direct control over the program during performance. In the scenario developed in the present study, *Mimi* generates music based entirely on the preparatory material, and is unaffected by the new human and machine improvisations during performance, so as to keep the source material manageable for both the improvisation system and the performer (see discussion in Section 4).

*Mimi*’s improvisation engine is inspired by that of the OMax system [1, 5, 2]. In OMax, Assayag, Dubnov et al. introduce the factor oracle approach in the context of machine improvisation. In OMax, off-line versions of the learning and generation algorithms are implemented in OpenMusic [11, 3, 4]. OpenMusic’s functional programming approach does not allow for efficient handling of real-time (or on-line) events. Consequently, OMax relies on Max/MSP [9] to handle on-line aspects such as real-time control, MIDI and audio acquisition, and rendering. OpenMusic and Max/MSP adopt similar visual metaphors (patches), but with different, and incompatible, semantics: as observed by Puckette in [14], Max patches contain dynamic process information, while OpenMusic patches contain static data. Therefore, communication between, and coordination of, the two subsystems in OMax requires the use of an interaction protocol, OpenSound Control [12]. *Mimi* is designed and implemented in a single formalism (SAI), which results in a simpler and more scalable system. For example, the integration of complex visualization functionalities occurred in a natural and seamless way, without compromising system interactivity.

In OMax, the improvising musician interacts with the system based on purely aural feedback, while another human operator controls the machine improvisation (the improvisation region in the oracle, instrumentation, etc.) through a visual interface, during performance. *Mimi* explores a different approach, with interaction specifically designed for, and under the sole control of, the improviser. Visual feedback of future and past musical material, and of high level structural information, provides timely cues for planning the improvisation during performance. The oracle on which the improvisation is based can be spontaneously generated, or pre-meditated, and carefully planned.

Factor oracle-based improvisation is based on a stochastic process in which the musical material generated is a recombination of musical material previously learned. Many improvisation systems make use of various probabilistic models to learn parameters from large bodies of musical material, then generate new material with similar statistical properties (akin to style). Examples of machine improvisation sys-

tems employing other probabilistic models include Pachet’s Continuator [13], Thom’s Band-OUT-of-the-Box (BoB) [16, 15], and Walker’s ImprovisationBuilder [19, 17, 18].

Musical interaction modalities during performance vary from system to system, and range from turn-taking dialogue to synchronous accompaniment, but in all cases, performer-machine interaction is based exclusively on aural feedback. Recently, the Haile/Pow humanoid-shaped robotic drummer [20] introduced gestural interaction for musical improvisation. *Mimi*’s visual feedback design does not aim to emulate human gestures, but rather, it seeks to explore different modes and mental spaces for communication in the interactive music creation process.

## 3. VISUAL INTERFACE

This section describes the key and novel component of the *Mimi* system, its visual interface, shown in Figure 2. The purpose of this interface is to provide the performer with instantaneous and continuous information on the state of the system. For human improvisation, in which context and planning are paramount, the relevant state of the system extends to the near future and recent past.

*Mimi*’s visuals adhere to well-known principles of design for usability and understandability [10], namely to provide a good conceptual model, visibility, feedback and natural mappings. The screen image is divided into two halves: the upper half of the visual display shows the interactions between the machine’s improvisations and the human performer’s improvisations, and the musical material to come, and recently passed, in real time. The lower half of the visual display documents the current state of the improvisation engine, with respect to the preparatory material.

The upper half, the *performance visualization* pane, displays a scrolling piano roll notation, showing past, present, and future notes, with a central bar marking the present. The notes float from right to left, and are sounded when they pass the center bar. Both machine and human improvisations are shown on the same panel, distinguished by color codes: blue for the machine, and red for the human. (Future human-originated notes are generally inaccessible, and therefore do not show on the panel.) The lower half, the *oracle visualization* pane, shows the source musical material for the factor oracle, provided in the preparatory stage.

As the initial material is created, the interface shows the incoming notes, in piano roll representation, scrolling from right to left in the performance pane, and collecting cumulatively from left to right in the oracle pane. In the performance phase, the system generates new musical material from the prepared material, with which the performer improvises. A red vertical line shows the present state of the oracle as a position in the source material. This cursor maps the traversal of the oracle links in real-time; its rate of non-sequential transitions reflects the amount of recombinations introduced by the oracle. This display presents a visual aid to the human improviser, helping the performer keep a mental note of the structural content and relative location (within the initial material) of the current machine improvisation. Meanwhile, the performance visualization acts as a musical map, allowing the performer to see the oncoming musical material and to plan ahead, as well as to visualize the interaction of the recent past improvisations between themselves and the machine.

When projected on a large screen in a performance, these

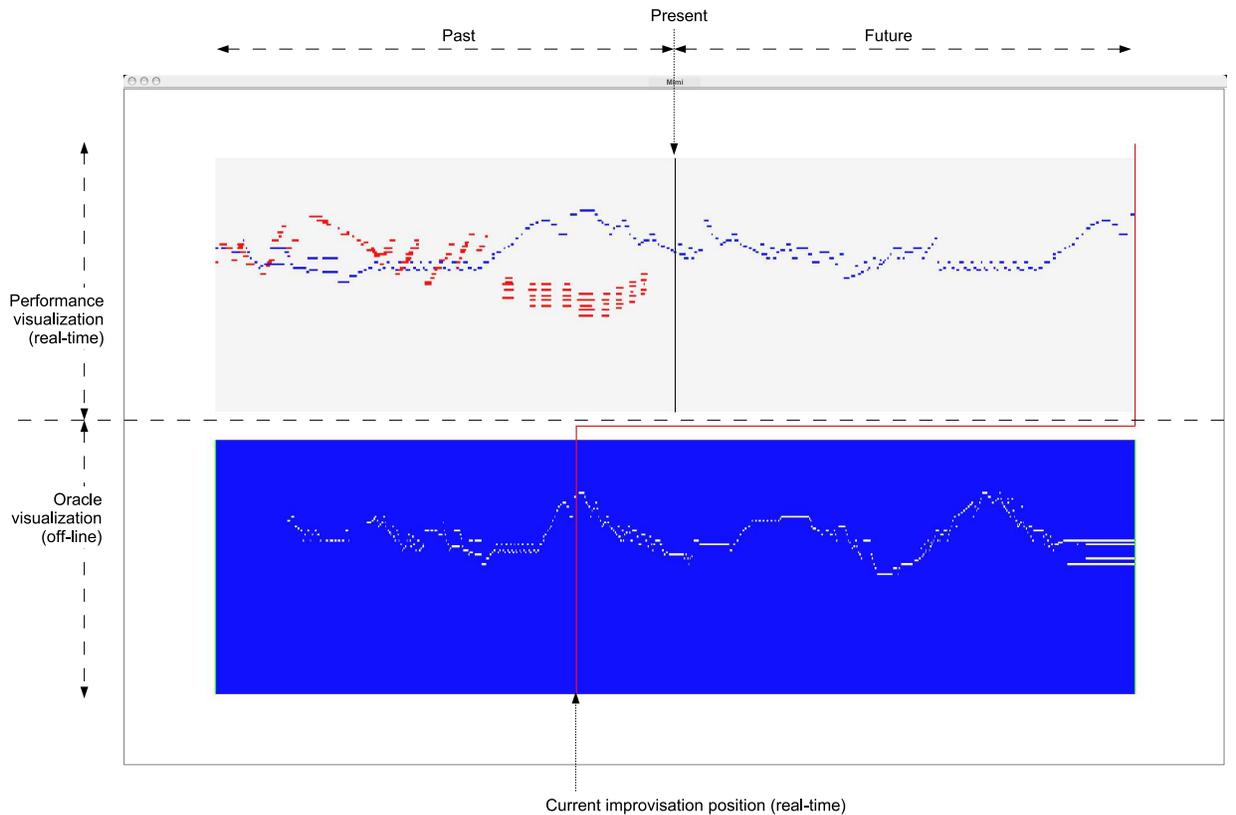


Figure 2: *Mimi*'s visual interface.

visuals may also assist the audience in perceiving higher level structures and the subtleties of improvisation strategies.

The piano roll conceptual model is familiar to musicians. The mapping of time along the horizontal axis is omnipresent in music notation schemes, and it accommodates both static (oracle) and dynamic (performance tracks) representations. Furthermore, the level of precision offered by the piano roll notation is particularly well suited to improvisation, which involves fast pattern identification and recognition. A more precise graphic notation, such as score notation, might overwhelm the less-experienced performer. On the contrary, the piano roll notation allows for the association of musical ideas with visual patterns.

#### 4. VISUALS IN IMPROVISATION

This section offers a discussion of the system and the motivation behind the design of its visual interface. *Mimi*'s design evolved over the course of quasi-weekly play sessions involving the three authors. In the early sessions, Thurmond exercises free license in experimenting with the latest version of the system, and learning its behavior in improvisation. As the sessions deepen in sophistication, the play sessions become more structured, and the laying down of preparatory material of each oracle, and the experimentation with its expressive scope, is preceded and succeeded by question and answer sessions to interrogate Thurmond on the interface design, and the oracle design and performer-machine improvisation process.

*Mimi* incorporates aspects of raw reflex typically associ-

ated with improvisation, and preparation and timing more closely affiliated with score-based reading. The visual feedback gives the performer foresight into the future (through the right-hand-side of the scrolling upper panel showing future musical material), hindsight into the past (through the left-hand-side of the scrolling upper panel), and the current state of the oracle within the preparatory material. These visual cues were incorporated after the first few sessions of free play, where it was determined that the performer needed more than simply a scrolling piano roll of the music's now and past.

Comparing the importance of the three types of information (future and past material, and oracle state), Thurmond reports that he pays attention to the future content part of the panel approximately 60% of the time, and divides the remaining 40% of the time between the past and the oracle state parts of the display. It is expected that the improviser would find the future musical content of imminent interest for planning purposes, to be able to foresee what the machine would be playing next ahead of time, and to have the time to prepare counter material in response.

Quoting Thurmond, "I'm looking at what's happening right then, and then I'm scanning ahead constantly. ... when I would see a certain kind of dissonant pattern coming up, I would know, because of the way I set up the oracle itself, what other dissonant intervals I could put with it. And then I would use clusters where I could see it coming to something I had set up that was open, and I would mix the cluster with it, and I would repeat the cluster. And sometimes, I could

see the silence, which is very important in the original oracle – to build in silence – then I could place that, and set up an ostinato, and it would hit it, and it was great. That was very exciting.”

What was surprising was the discovery that the past information was also important. As Thurmond puts it, it is important to look at what’s gone after “because it gives me an idea of what I want to do in the future. ... having a history is very important because if I do something that is not quite right, in improvisation if you make a mistake, then you don’t jump away from it, you stay with it to make it not a mistake any more. ... It becomes a feature ... then I bring it back again.” The statement resonates with John Cage’s quote, “The idea of mistake is beside the point, for once anything happens it authentically is.” When the unexpected happens, to re-create it again, one has to recognize how it happened, and in what context, so as to be able to repeat a particular pattern in the future. This is where the visual display of the past, of the machine-generated context (in blue) and the human response (in red) comes into play.

Thurmond uses the oracle state panel in the lower half of the display, to contextualize the present musical content and state of the machine improvisation. “I wanted to see what [*Mimi*] was thinking. ... it allowed me a lot more freedom than trying to remember just exactly what kind of pattern I set up. So the structure that I set up, I am constantly reminded of the structure.” The oracle’s state thus provides the performer with a quick reference for the structure of the present improvisation material.

## 5. CONCLUSION

This paper described the design of *Mimi*, a musical improvisation system that explores the potential and powerful impact of visual feedback in performer-machine interaction. In only a few short months, it has already proven itself to be an invaluable platform through which to investigate the mental models necessary for successful improvisation. Future research will explore the improvisation strategies through the defining of key principles in setting up effective oracles in *Mimi*’s preparatory stage, and its application to the teaching of improvisation.

## 6. ACKNOWLEDGMENTS

The authors thank Gerard Assayag for his valuable discussions on, and insights into, the use of factor oracles for machine improvisation in OMax, and Katherine deSousa (USC Women in Science and Engineering Undergraduate Research Fellow) for her help in video taping and transcribing the performer-machine improvisation sessions.

## 7. REFERENCES

- [1] G. Assayag, G. Bloch, M. Chemillier, A. Cont, and S. Dubnov. Omax brothers : a dynamic topology of agents for improvization learning. In *Proc. ACM Workshop on Music and Audio Computing*, Santa Barbara, CA, USA, October 2006.
- [2] G. Assayag and S. Dubnov. Using factor oracles for machine improvisation. *Soft Computing*, 8:1–7, 2004.
- [3] G. Assayag, C. Rueda, M. Laurson, C. Agon, and O. Delerue. Computer assisted composition at IRCAM: PatchWork & OpenMusic. *Computer Music Journal*, 23(3), 1999.
- [4] J. Bresson, C. Agon, and G. Assayag. OpenMusic 5: A cross-platform release of the computer-assisted composition environment. In *Proc. 10th Brazilian Symposium on Computer Music*, Belo Horizonte, Brazil, 2005.
- [5] S. Dubnov and G. Assayag. Improvisation planning and jam session design using concepts of sequence variation and flow experience. In *Proc. International Conference on Sound and Music Computing*, Salerno, Italy, November 2005.
- [6] A. R. François. A hybrid architectural style for distributed parallel processing of generic data streams. In *Proc. International Conference on Software Engineering*, pages 367–376, Edinburgh, Scotland, UK, May 2004.
- [7] A. R. François and E. Chew. An architectural framework for interactive music systems. In *Proc. International Conference on New Interfaces for Musical Expression*, Paris, France, June 2006.
- [8] G. Lewis. Too many notes: Computers, complexity and culture in voyager. *Leonardo Music Journal*, 10:pp. 33–39, 2000.
- [9] Max/MSP. [www.cycling74.com](http://www.cycling74.com).
- [10] D. A. Norman. *The Design of Everyday Things*. Basic Books, 2002.
- [11] OpenMusic. [recherche.ircam.fr/equipes/repmaus/OpenMusic](http://recherche.ircam.fr/equipes/repmaus/OpenMusic).
- [12] OpenSound Control. [www.cnmat.berkeley.edu/OpenSoundControl](http://www.cnmat.berkeley.edu/OpenSoundControl).
- [13] F. Pachet. The continuator: Musical interaction with style. *Journal of New Music Research*, 32(3):333–341, 2003.
- [14] M. S. Puckette. A divide between ‘compositional’ and ‘performative’ aspects of Pd. In *Proc. First International Pd Convention*, Graz, Austria, 2004.
- [15] B. Thom. Bob: an interactive improvisational companion. In *Proc. International Conference on Autonomous Agents (Agents-2000)*, Barcelona, Spain, 2000.
- [16] B. Thom. Interactive improvisational music companionship: A user-modeling approach. *The User Modeling and User-Adapted Interaction Journal; Special Issue on User Modeling and Intelligent Agents*, Spring 2003.
- [17] W. Walker and B. Belet. Applying ImprovisationBuilder to interactive composition with midi piano. In *Proc. International Computer Music Conference*, Hong Kong, China, 1996.
- [18] W. Walker, K. Hebel, S. Martirano, and C. Scaletti. ImprovisationBuilder: Improvisation as conversation. In *Proc. International Computer Music Conference*, San Jose, CA, USA, 1992.
- [19] W. F. Walker. A computer participant in musical improvisation. In *Proc. Human Factors in Computing Systems (CHI)*, Atlanta, GA, USA, March 1997.
- [20] G. Weinberg and S. Driscoll. Robot-human interaction with an anthropomorphic percussionist. In *Proc. Human Factors in Computing Systems (CHI)*, Montreal, Québec, Canada, April 2006.