

# A Personal Chronology of Audiovisual Systems Research

Golan Levin

Carnegie Mellon University  
College of Fine Arts, CFA-300  
5000 Forbes Avenue  
Pittsburgh, PA, 15213 USA  
+1.412.268.2000

golan@andrew.cmu.edu

## ABSTRACT

In this invited lecture, I present an informal overview of seven years' research into the design of real-time systems for the creation, manipulation and performance of simultaneous image and sound. This research explores the intersection of abstract communication and interactivity, as part of a more general inquiry into the formal languages of the responsive medium, and of nonverbal communications protocols in cybernetic systems. I present a combination of live demonstrations and video documentations in order to illustrate the various systems, reveal some common threads, and propose some design desiderata.

## Keywords

Audiovisual performance, interactive art, new media, live cinema, dynamic abstraction.

## 1. INTRODUCTION

For the last several years, my collaborators and I have researched the design of interactive systems for the creation and live performance of abstract image and synthetic sound. These systems have not, generally speaking, investigated mappings between sound and image *per se*, but rather the ways in which audiovisuals can be linked to expressive human gestures such as hand movements and speech. Ideally, these systems are intended to encourage people to engage in flow states of pure experience.

## 2. SYSTEMS

In my talk, I will present demonstrations and documentations of a variety of interactive projects, including the following:

### 2.1 The Audiovisual Environment Suite

The *Audiovisual Environment Suite* [AVES] (2000) is a set of set of six mouse-driven software artworks which allow people to gesturally create and perform abstract animation and synthetic sound in real-time. Each environment is an experimental attempt to design an interface which is supple and easy to learn, yet can

also yield interesting, infinitely variable and personally expressive performances in both the visual and aural domains. The systems are based on the metaphor of an "audiovisual substance" which can be created, manipulated, and deleted in a painterly, non-diagrammatic image space [1].

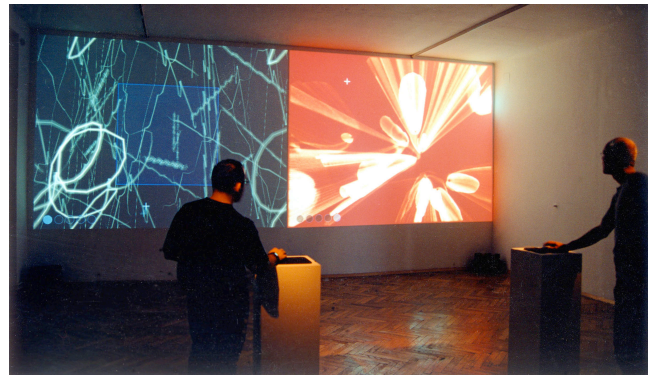


Figure 1. The *Audiovisual Environment Suite*.

### 2.2 *Dialtones: A Telesymphony*

*Dialtones: A Telesymphony* (2001, with Gregory Shakar, Scott Gibbons, Yasmin Sohrawardy, et al.) is a performance whose sounds are wholly produced through the carefully choreographed ringing of the audience's own mobile phones. Before the concert, participants register their mobile phone numbers at a series of web terminals; in exchange, new ringtones are automatically transmitted to their phones, and their seating assignment tickets are generated. During the concert, the audience's phones are dialed up by live performers, using custom software which permits as many as 60 phones to ring simultaneously. Because the exact location and tone of each participant's phone is known in advance, *Dialtones* offers a wide range of interesting sonic phenomena, such as waves of polyphony which cascade across the audience. Additionally, owing to a special projection system, each member of the audience becomes a uniquely addressable audiovisual pixel [2].

### 2.3 *RE:Mark and Hidden Worlds*

*RE:Mark and Hidden Worlds* (2002, with Zachary Lieberman) are a pair of interactive audiovisual installations in which participants are able to "see" each others' voices, made visible in the form of animated graphic figurations that appear to emerge from the participants' mouths. *RE:Mark* uses simple computer vision and projection technologies to create the fiction that speech casts

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, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

*Nime'05*, May 26-28, 2005, Vancouver, BC, Canada.  
Copyright remains with the author(s).

visible shadows. *Hidden Worlds*, on the other hand, is an augmented reality, in which special glasses superimpose 3D graphics into the real world; when one of the users speaks, colorful abstract forms appear to emerge from her mouth [3, 4].

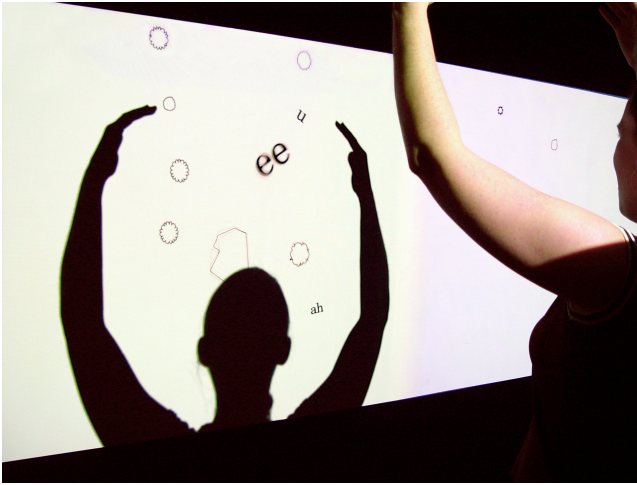


Figure 2. *RE:Mark*.

## 2.4 *Messa di Voce*.

*Messa di Voce* (2003, with Zachary Lieberman, Jaap Blonk, and Joan La Barbara) is a performance in which the speech and songs produced by two vocalists are augmented in real-time by custom interactive visualizations. Our software transforms vocal nuances into correspondingly complex, subtly differentiated and highly expressive graphics. These visuals not only depict the singers' voices, but in some circumstances also serve as controls for their acoustic playback [4, 5].

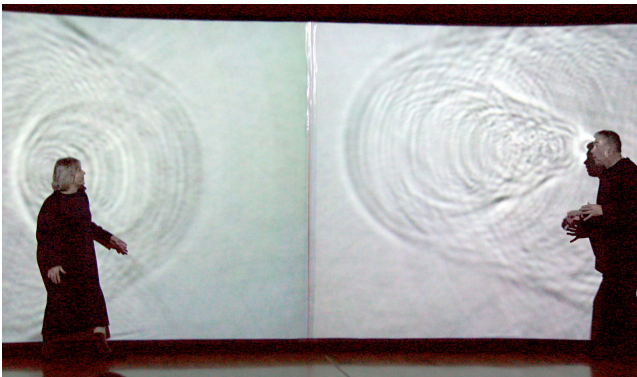


Figure 3. Joan La Barbara and Jaap Blonk perform real-time speech visualizations in *Messa di Voce*.

## 2.5 *The Manual Input Sessions*.

*The Manual Input Sessions* (2004, with Zachary Lieberman) is a series of audiovisual vignettes which probe the expressive possibilities of hand gestures and finger movements. This concert is performed on a combination of custom interactive software, analog overhead projectors and digital computer video projectors. The analog and digital projectors are aligned such that their projections overlap, resulting in an unusual quality of hybridized,

dynamic light. During the performance, a vision system analyses the silhouettes of the performers' hands as they move across the glass tops of the overhead projectors. In response, our software synthesizes graphics and sounds that are tightly coupled to the forms and movements of the performers' actions. The graphics are co-projected over the analog shadows, resulting in an unusual form of augmented-reality shadow play [4, 6].



Figure 4. *The Manual Input Sessions*.

## 3. ACKNOWLEDGMENTS

These works would not have been possible without the loving support of my wife, Andrea. I am also greatly indebted to the many intrepid collaborators who have shared in the pain and pleasure of developing these systems with me, notably including Zachary Lieberman, Greg Shakar, and Scott Gibbons, as well as Jaap Blonk, Cassidy Curtis, Jonathan Feinberg, Joan La Barbara, Scott Snibbe, Yasmin Sohrawardy, Martin Wattenberg, and Shelly Wynecoop.

## 4. REFERENCES

- [1] Levin, G. "Painterly Interfaces for Audiovisual Performance." M.S. Thesis, MIT Media Laboratory, August 2000. <http://acg.media.mit.edu/people/golan/thesis/>.
- [2] Levin, G, et al. *Dialtones*. <http://www.telesymphony.com/>.
- [3] Levin, G. *Flong*. <http://www.flong.com/>.
- [4] Levin, G. and Lieberman, Z. *Tmema*. <http://www.tmema.org/>.
- [5] Levin, G. and Lieberman, Z. "In-Situ Speech Visualization in Real-Time Interactive Installation and Performance." *Proc. 3rd International Symposium on Non-Photorealistic Animation and Rendering*, Annecy, France, 2004.
- [6] Levin, G. and Lieberman, Z. "Sounds from Shapes: Audiovisual Performance with Hand Silhouette Contours in *The Manual Input Sessions*." *Proceedings of NIME '05*, Vancouver, BC, Canada. May 26-28, 2005.
- [7] Snibbe, S. and Levin, G. "Instruments for Dynamic Abstraction." *Proceedings of the First Annual Conference on Non-Photorealistic Animation and Rendering*, Annecy, France, 2000.