Ozone: Continuous State-based Media Choreography System for Live Performance

Sha Xin Wei
shaxinwei@gmail.com

Michael Fortin
michael.fortin@gmail.com

Navid Navab
navid.nav@gmail.com

Tim Sutton
timsutton@fastmail.fm

Topological Media Lab
Concordia University
Montreal, Qébec, Canada

ABSTRACT

This paper describes Ozone, a new *media choreography* system based on layered, continuous physical models, designed for building a diverse range of interactive spaces that coordinate arbitrary streams of video and audio synthesized in realtime response to continuous, concurrent activity by people in a live event. We aim to build rich responsive spaces that sustain the *free improvisation of collectively or individually meaningful, non-linguistic gesture*. Ozone provides an expressive way to compose the potential "landscape" of an event evolving according to the designer's intent as well as contingent activity. A potential-energy engine evolves superposed states over simplicial complexes modeling the topological space of metaphorical states.

Categories and Subject Descriptors: J.5 [Arts & Humanities]: Performing arts, C.3 [Special-purpose & Application-based Systems]: Real-time and embedded systems, I.4 [Image Processing & Computer Vision]: Applications, D.2 [Programming Environments]: Interactive environments, H.5 [User Interfaces]: Input devices and strategies

General Terms: Design, Experimentation, Human Factors, Performance, Theory

Keywords:responsive environment; media choreography; immersive space; responsive media; dynamical system; physics model; simplicial complex; gesture-based media; realtime video; realtime sound; wearable computing, design.

1. INTRODUCTION AND APPROACH

Typical interactive installations or environments exhibit a behavior that, however complex, does not vary in quality over time, or is governed by a branching logic that feels artificial or sparse compared with improvised, non-schema-bound, playful activity in a live event. Making media environments rich usually comes at the expense of custom-coding singular systems due to a lack of a human-legible representations of event dynamics that is not wired down to a specific event structure.

MM'10, October 25–29, 2010, Firenze, Italy.

The Ozone *media choreography system* is a software system for choreographing media in concert with people's activities in a responsive environment. It maps sensor data into fields of time-based media such as realtime video and sound derived from continuous states evolving over a simplicial complex defined on metaphorical labels on possible states of an event imagined by a designer. Ozone's physics engine for continuous states allows composers to intuitively design a "landscape" of possible states that evolve in response to the players' activities. The state evolution system generalizes from a finite-state machine to *continuous* evolution of *superposable* states with reproducible kinematics. Ozone is as rich as a custom built, singular interactive environment, yet is a general composition system that can yield a great variety of installation-events (q.v. Section 5).

Ozone's quasi-physics model drives dense fields of video and sound that can be intuitively manipulated in real-time by the *player*, a visitor in the environment, like swimming or watering a garden.

Our central technical interaction design challenge is how to make a richly playable responsive media space that has no pre-assigned interface objects nor pre-specified gesture/action sequences. A typical "interactive" installation has a behavior that, however rich in its basic dynamics, basically does not meaningfully change its state: an eternal thunderstorm of particles, for example. Ozone allows the designer to compose a responsive media environment whose behavior qualitatively changes in a palpable way according to both the composer's design and contingent activity. As the players interact with the projected sound and imagery over time, they can *invent* continuous ranges of gestures to meaningfully shape the media. Thus we do not require pre-fabricated user interface objects or pre-specified gestures but instead allow players to construct manipulable objects out of the media textures themselves.

As an event composition system, the challenge is to support designers who wish to dispense with locally one-dimensional narrative structure and build instead a dynamical ecology of image and sound. Since designers very carefully craft media textures according to musical and visual symbolism, the system must support the control of richly structured transformations rather than just arbitrary textural transformations on media like "video wallpaper". At the same time, the dynamics need to be tightly coupled to the gestures and movement of the players in the room, as well as to the internal state of the system. Traditional timeline-based narrative scripts and database representations do not support such complex dynamics. Representations in conventional procedural languages are too low-level to efficiently capture the complex semantics of the artists' metaphorical designs. We pitch the Ozone's representational structure at an intermediate level between the low-level response to sensor-data plus their statistical derivatives, and the high-level semantics of the artists' metaphorical talk.

Figure 1: Player in TGarden, Ars Electronica 2001 *A.M. Korpi.*

Ozone's pseudo-physics and high-level evolution of the "climate" of the environment are articulated by representations (Figure 3) that are legible to dramaturgical and aesthetic designers who are not programmers. Since the pseudo-physics *constitutes* the experience of a responsive environment, the design of the pseudo-physics constitutes the explicit expressive design of the composer of such a responsive space. (For a deeper discussion of such performative and dramaturgical requirements, see [23].)

For example, a player can be described as a soloist or as part of a group. But rather than simply oscillate between two discrete player states Solo and Group, we generalize this aspect of the player to a continuous, *multi-dimensional* range of "groupness", a mixture between Solo and Group states. Thus, this approach provides a level of interpretation and indirection between the sensors and the rest of the system; the designers can iteratively shape a richly evolving experience without binding to a particular fixed set of physical, observable behaviors. For example, once groupness is defined as a concept based on the two metaphorical components Solo and Group, the designers can begin using it in the media synthesis engine. A player's groupness could be indexed to the distance from that player to the barycenter of all the players in the room, *or* it could be indexed to a measure of synchrony between that player's accelerometer data with the other players' accelerometer data. The final choice can be made in the rehearsal phases of the production of a responsive environment, with designers and players working together in a particular performance scenario.

We first situate the state engine relative to prior work in Section 2. We describe our methods in Section 3, the Ozone system architecture, and Section 4, the explicit model for the continuous state dynamics. In Section 5, we describe the results of the Ozone system, as applications. In Section 6, we discuss how designers can work with Ozone. We close with future research directions.

2. RELATED WORK

One of the pioneers of a dynamical systems approach to a media environment was Gordon Pask, whose interest extended to second order cybernetic systems after Heinz van Foerster. Although his work was primarily in theories of learning and conversation, early on, Pask created with Robin McKinnon-Wood an electromechanical system called MusiColour which modulated colored lights in response to live music. When MusiColour habituated to a pattern, it would offer *adversarial* patterns increasingly aggressively until the musicians shifted to another pattern [7].

Myron Krueger [10] pioneered augmented reality installations in a related spirit, but a different perspective. He proposed that the proper medium of a responsive space is not image or sound, but the responsivity itself. Though this blurs phenomenological categories, it's a useful motto for the architecture of responsive media spaces.

The idea of computational steering by one or several users was introduced in scientific simulation ([13], [20]) when real-time control of complex dynamical systems became computationally feasible. Those applications were directed to visualizing natural processes rather than augmented physical performance environments, but they pointed a way to contemporary applications.

People have used dynamic physical simulation for such applications as graph layout ([17], [8]). Our approach attempts in a parallel manner to imbue general projected visual and audio environments with more usable palpability[1]. But instead of relying on non-deterministic state machines to deal with state transitions when the input relies on recognition of noisy input, as [8] does, an Ozone system's behavior evolves according to an intrinsic physics, which lends it palpability and intuitability[2].

The response logic in interactive spaces such as the "Kids Room" [1] or "It/I" [16] – is based on narrative scripts and physical or virtual actors, whereas the Ozone system is built to support ecological models of evolution. However, our continuous state topology can represent decision-tree logics such as branching narratives for games as a discrete special case. Ozone accommodates not only rehearsed performers but also people informally improvising gesture using casual movements and gestures in everyday situations.

Outside the laboratory, conventional show control systems function robustly and accommodate the full spectrum of device control and communication protocols (DMX, MIDI, RS232, TCP etc.). However, their sub-narrative device control logic is too primitive to accommodate the design semantics and dynamics of media choreography that are called for in a performatively rich media space like TGarden.

Laboratory built media stream processing systems such as those used in the Intelligent Stage [15] are subsumed by professional, robust, and user community-supported composition environments like Max, hence Ozone's reliance on a modular architecture that has accommodated wholesale changes in hardware platforms and software implementation environments. From a usability perspective, however, it is key to focus composer or choreographer's energy not on configuring the physical devices or computational units, but on structuring the *event*. Ozone differs from most stream management systems ([15], [14], [5], [12]) in that the focus of attention is on the event, and in fact, on the *potential space* of events and evolution of state in response to activity and design.

[3] and [19] survey related classes of responsive environments or interactive spaces.

[6] has speculatively compared standard interfaces with musical instruments like the violin. In such spirit, motivated by our collaboration with designers of theatrical and interactive spaces, we decided to start by looking to designers whose expertise lay with gesture-based electronic musical instruments[3] and with experimental garment design.

[11] and [9] have used simplicial complexes to model a configuration space for the purposes of constraining freehand tugging or morphing to meaningful configurations in a subset of possible diagram structures. This powerful approach depends on building a representation at a fairly abstract level. In contrast, our topologi-

[1] See hopskip palpability example: http://vimeo.com/10820127

[2] The woman discovers *with no hints* a way to play with the "jewelry" as a sonic controller: http://vimeo.com/10820175

[3] Such as the Input Devices and Music Interaction Laboratory at McGill University, www.idmil.org/; and the Studio for Electro-Instrumental Music, http://www.steim.nl

cal model for state space is built directly in the metaphorical language used by the designers. The Ozone system deals with general time-based media projected into a physical space, allowing for autonomous processes with parameters that allow continuous-time steering driven by live human gestures and movements.

The prototypical responsive media environment motivating Ozone is the TGarden, designed in 1997 as a responsive media playspace and realized in 2001 by the Sponge and FoAM collectives, with costumes and wireless sensors, IR tracking, and realtime sound and video synthesis to support improvised solo and group play. The TGarden was tested under harsh performance conditions in 2001 at the Ars Electronica and V2 by the Sponge and FoAM, which provided valuable feedback on the system design.

Beyond the motivating TG2001 and txOom applications, we have exercised different components of the Ozone system in the following installations and events: MeteorShower 2006 (sound, video from particle systems driven by video-based motion-tracking); Cosmicomics 2007 (public installations with responsive sound and video, and linear states), Ouija 2007 (dance and movement with responsive sound and video); Remedios Terrarium 2008 (public exhibit of complex self-interacting sound, video-tracking, and state); and Frankenstein's Ghosts 2008-2010 (performance research augmenting the *Blue Riders* chamber music ensemble and dancers' movement with responsive sound). [21, 22] detail the artistic and conceptual impacts of the environments built using our system.

One could create media responses to human activity based on explicit deterministic algorithms or on fixed scripts. However, the designers of the responsive media spaces that we are supporting have no deterministic, algorithmic methods that would express the phenomenological richness and cultural symbolism of the experiences that they wish to sustain. Alternatively, one could try to accomplish all this with a system oriented toward gesture recognition, but such a system would be much too rigid for our purposes, in which we desire to provide a substrate for intuitive understanding, navigation and manipulation of a rich media environment without resorting to (grammatical) language or other explicit formal representations.

Finally, Ozone's energetic model differs from statistical methods such as Markov random fields or hidden Markov models in two fundamental ways. First, Ozone is modeled over a simplicial complex rather than a graph, and consequently sustains much more general spaces of behavior. Second, our energetic model has an interpretation in terms of classical mechanics, rather than a statistical one, and our player state moves in a continuous[4] and deterministically, rather than discretely and randomly across the state space.

3. OZONE

Ozone has the following design features: (D1) The designer, actor, and spectator may be the same body, implying that we focus on first-person experience; (D2) The primary modes of interaction are not based on (isomorphs of) linguistic patterns, but on continuous fields of matter and media; (D3) The participants are always in a common physical place, placing a very high bar on perceptual density and zero effective latency; (D4) The designer composes not specific event sequences but *meta*-events, or substrate potential-energy fields that condition continuous sets of possible events; (D5) We design for continuous experience with the density of everyday settings.

Technically, the choreography system is also designed to: (D6) directly accommodate high-level designer semantics for interpreting or responding to player activity; (D7) incorporate arbitrary continuous as well as discrete evolution of state; (D8) support low latency responsiveness to sensor data; and (D9) support realtime synthesis of structured and rich continuous video and audio.

[4]"Continuous" is meant here in the sense of a topologically continuous space, modeled by finite-difference approximations.

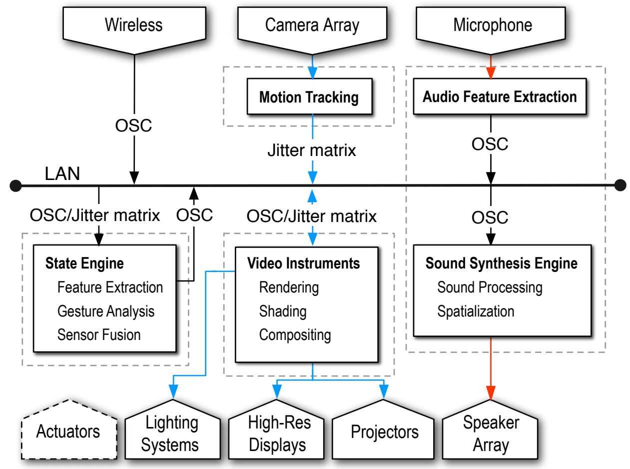

Figure 2: Ozone architecture: components and data paths. *Diagram by Morgan Sutherland*

In subsequent sections, we will mark where the engineering and mathematical models meet these desiderata.

3.1 Hardware and software platforms

The architecture of the current Ozone system is illustrated in Figure 2. Our realtime system is written as a set of components typically written in C, C++, Max / MSP / Jitter, each on a separate host: (1) a data routing, statistics, gesture tracking and media choreography engine written in Max[5], (2) a sound synthesis engine written in Max with externals such as a granular synthesis and DSP analysis, and (3) video effects engine written in Max[6]. We chose to implement our media synthesis instruments in realtime processing environments as Max and SuperCollider since these are among the most expressive, popular, and professionally maintained systems currently used as a *lingua franca* by media artists and musicians. Thus we efficiently incorporate rich musical and visual imagery and dynamics without writing special purpose effects from scratch that would be restricted to laboratory demos.

In addition to these core system components, we have used auxiliary devices to sense activity. We have used CrossBow's low overhead TinyOS wireless sensor platforms, with a sensor-board containing a magnetometer, accelerometer, photocell, sound meter, and analog and digital inputs for additional sensors. While convenient and powerful as development environments, these sensor platforms tend to be fragile for moving bodies. We anticipate migrating computation between body-borne or fixed hosts as the trade-offs between platforms, bandwidths, and power supply evolve. Currently, where possible we use IR camera-based computer vision to track movement and form to avoid instrumenting the participants.

Berkeley CNMAT's Open Sound Control (OSC) protocol[4] forms the basis for communication among the various components of the Ozone system. The OSC protocol has several advantages, such as naming of individual resources, ease of use, availability in various development environments, time-tags on data packets, and decoupling from network transport protocol (UDP, TCP-IP, etc.).

[5]http://cycling74.com

[6]Preceding instruments were written in Supercollider http://www.audiosynth.com by Joel Ryan, and in NATO http:// www.m9ndfukc.org/ korporat/ nato.0+55+3d.html

With OSC we can treat each host uniformly whether it is a sensor platform on a moving body or a fixed computer.

3.2 Feature Extraction

Given a set of sensors, one of the key problems to be solved by any interactive media system is how to reduce multiple time-series of sensor data to extract useful features. The Ozone system performs relatively simple statistical reductions e.g. on time- or sample-based moving window averages or $W^{p,q}$ norms, and aggregations across multiple sensor streams and multiple players. An example is the total angular momentum of the players' motions about the center of the floor.

One could perform the sensor data reduction by an automatic supervised or unsupervised procedure, adapting, for example, an independent component analysis relying on some appropriate entropic measure. However, in order to achieve meaningful performative, experiential impact, it is important that the reduction of sensor data, especially that part which is mapped responsively to the audiovisual synthesizers, be understandable by the players and designers. We believe that a good way to achieve this is to allow the designers to manually specify the sensor feature mapping.

3.2.1 Video analysis

Most of our video analysis is done live (D8, D9) using standard computer vision tools such as optical flow measures or morphological filters from IAMAS' cv.jit library. We also built moving window background thresholding and adaptive motion extraction extensions. We treat video input as an array of pixel-sensors $O(10^5)$ denser than our amorphous networks of physical sensors.

3.2.2 Audio signal analysis

Continuously running video analysis (motion detection and basic computer vision features) is depended upon by multiple nodes within Ozone to evaluate existing and generate new behaviors, and similarly acoustic information pertaining to a space should be available (D8, D9).

Audio analysis modules are a combination of common signal analysis algorithms (spectral centroid, Bark scale frequency amplitudes, envelope following) and derivatives of these features that can be used compositionally, both as controls by which other system components behave, and to themselves inject biased dynamics into other media elements in the system. An example of a derivative feature would be a two-stage cascaded leaky integrator, which supplies two different "rates" of activity of a parameter such as "activity," based on a conditioned spectral centroid analysis of a table-mounted boundary microphone. The dynamics of these varying rates in addition to raw, fast sensor data provide a more rich data set to the compositional state system than simplistic one-to-one linear relationships of sound to other media.

In improvisatory events, we cannot assume our sensor data conform to an *a priori* model. Indeed, such approaches to signal analysis can behave unexpectedly due to unpredictable activity, changing environmental conditions (lighting, acoustics), noise at various stages in the system, and feedback from the media systems themselves. The behavior of the algorithms implemented in a given configuration can fluctuate, as we cannot in every case predict a standard input data range or dynamic.

3.3 Time Scales

While the media choreography engine runs in realtime, the architecture of the system has the side effect of causing the system to appear to run at three different time scales. This allows the system to simultaneously sustain a sense of tangibility based on fine-grain temporal response based on features derived from sensor data (D8), as well as of global evolution of metaphorical state.

In practice, we found three scales of temporal dynamics to be meaningful to the designers: *micro*: $O(10^{-2})$ sec sensor data (e.g., de-noising at the sensor PIC), *meso*: $O(1) - O(10)$ sec: designer-specified gestural grain (e.g., the rate at which perceptible changes in the value of a fundamental state like *Solo* occur), and *macro*: $O(10^2)$ sec "narrative" state.

The "hopskip" video shows how in practice this response allows a player to jump up and down and see his accelerometer's physical ballistics mapped to a graphic response (in the simple case shown, parametrized texture-mapping from a pre-fabricated video). The tight coupling gave the player the impression of *palpability* – that he was jumping on an elastic membrane, encouraging him to jump near the apparent characteristic frequency of the system.

3.4 Media Synthesis

This paper focuses on the physics-model based core of the media choreography system, but equally important are the *realtime* sound and image synthesis instruments written in Max (D9). We designed the instruments to be controlled by data streams such as raw and "cooked" sensor data as well as the state of the event (D8).

We parameterize our media instruments on physics-based state models so they can exhibit intuitable behavior according to event state while maintaining rich dynamic aesthetics (D2). The advantage of parameterizing the media instruments by (vector) state is that we do not have to program the combinatorially complex logics of how they vary according to every possible sequence of activity ahead of time. Consider a user-directed (D1) scene change between day and night. If the day and night are discrete states with a uni-dimensional transition arc between them, it would be hard to continuously, multi-dimensionally vary the saturation of color as well as the acoustic response due to humidity during sunset or sunrise. In fact, the user could improvisatorily vary the transition in a multi-dimensional mixture between night and day.

3.4.1 Video synthesis

Over the past decade, we have created several generations of realtime video processing abstractions in NATO, and now in Jitter. Our physics-based *calligraphic video* instruments [7] are built to respond to live video streams with negligible latency (D3).

Using the Max/Jitter framework, we can implement processing instruments on multi-dimensional dense lattices at video rates, streamed between local networks of hosts to parallelize the computation. An example of a visual instrument is MeteorShower (2006) which integrates a set of particles across a gravity field due to a 2D lattice of attractors. Mapping the attractors' masses and locations into a grid yields a computation speed that scales to the number of grid cells but not the number of people or moving objects in camera view (D2). Over the past three years, we have built an extensive library of CPU and GPU based realtime video instruments. We describe the computational physics and parallelization strategies on different hardware in [18].

3.4.2 Audio synthesis

In addition to a suite of custom-crafted instruments such as a wind generator and the polyphonic vocal synthesis engine used in the Meteor Shower [25] and Cosmicomics [24] installations and the Ouija movement experiment[8], Ozone's sound synthesis is built with a granular synthesis framework (munger1~ [2]). Granular synthesis has been the preferred method of sound synthesis because

[7] Calligraphic video: http://vimeo.com/10826801

[8] Ouija Calligraphy table: http://vimeo.com/10823874

of its wide range of timbral possibilities, both in terms of seed sound material (any PCM audio file) and multiple levels of parameter modulation. In fact, we consider any unique configuration (i.e. "preset") of this system to be its own instrument, with its own set of meta-parameters exposed to hide the preset details. Such parameters should be meaningful and non-technical (ie. "sludginess"). Considerable care to render the mappings in human-readable form allows the designer to rapidly audition very different sonic behaviors.

A challenge in working with different types of fast sensor input is that the sonic dynamics are often tightly coupled to those particular to a given sensor, which is a challenge for calibration and normalization of parameters throughout the system. A granular instrument's meta-parameters are designed to operate within a 0.0 to 1.0 range, but our instruments accept parameters outside this nominal range to permit other applications to drive them to unanticipated expressive results. This motivates our interest in simulating analog physical materials[18] using unconditionally stable algorithms where violent perturbations can yield qualitatively large yet steerable effects (D2, D3).

This decision comes from early experiments involving mappings of various (and variously-conditioned) channels of sensor and video analysis input to 6-10 low-level parameters, wherein the unpredictability of input dynamics further enriches the process of composing responsive media. If input constraints are desirable, for either technical or compositional reasons, they are programmed into a specific mapping as opposed to being built into the framework from which instruments are designed.

3.4.3 Spatial audio synthesis

Spatialization is the multi-layered process of controlling sound diffusion from speakers, and localizing sound sources within a reproduced or synthesized sound environment with specific spatial acoustic characteristics. Our spatialization toolkit bridges two core DSP macro-units composed of (1) algorithms for the encoding / decoding of the spatial transformations of sound-fields using Ambisonics (ICST externals [26]) and (2) room-effect synthesis engines for modeling and transforming spatial acoustics (SPAT [27]). Further, utility modules facilitate the control over the movement of sources and their perceptual attributes using a combination of algorithms, gestural controllers, and user definable mappings from our sensor data and state-engine. Our goal is to maximize intuitive thinking about sound and create spatial poetry by developing direct contact with continuously evolving spatialization parameters that are physical metaphors instead of technological abstractions.

Our spatialization toolkit has been optimized to incorporate many paradigms when needed such as: Vector Base Amplitude Panning (VBAP), Binaural, Distance-Based Amplitude Panning (DBAP), Surround, and etc. However, with testing we have gravitated towards Higher Order Ambisonics (HOA) for its flexibility, adaptability, and effective control over continuous transformations. HOA allows us to maintain dynamic and continuously evolving soundfields that could directly relate to the behavior of our state engine. When decoding Ambisonic signals, the directional response pattern of the generated soundfield can be narrowed or widened by adjusting the weighting of the different order components.

We are developing and exploring continuous transformations that link the spatial foreground and background, ambient soundscapes and dominant sound objects, and localizable crystals and immersive sound clouds. Integrating these software techniques with compositional and theatrical processes, we have for example been able to focus on the actual acoustic sound of sand as emitting from a stage and gently turn it into a massive and immersive wind that slowly transforms into a sharp fire-crackling stone that moves around the space (D2).

Ambisonics theory is partly based on the decomposition of a soundfield into spherical harmonics. The system will first encode sound sources and their induced spatial transformations into a flexible format, (HOA) that is later decoded in order for a soundfield to be projected as effectively as possible through the available speaker setup. The HOA signals are derived from a Fourier series on a spherical surface and together represent a three-dimensional soundfield. The decoding layer will then drive signals for all of the available speakers by combining the HOA signals depending on the actual position of the speakers in relation to the an imaginary sphere. In this way the decoder creates an accurate soundfield when possible and otherwise uses an approximation that takes the psychoacoustics of spatial sound into account. Therefore, unlike most surround systems, the speaker positions are not presumed or special and sound arriving from all direction are treated equally. One of the qualitative benefits of HOA is that the sound does not seem to come from the speakers but it is simply there and also the stability of the image depends less on the listeners position (D3). Moreover, if the number and placement of speakers change during the course of a project or event, the spatial thinking in the sound design will stay the same and remain easily reproducible in different spaces and systems. These characteristics make our spatialization software quickly adaptable to any possible spatialization scenario and speaker configuration, be it a theatrical production or a public installation.

4. CONTINUOUS STATE DYNAMICS

In this section, we explicitly describe the potential energy based dynamical system which models the meaningful states of configurations of people, activity, parameterized by live sensor data. This model compactly captures the metaphorical ontology and dynamical response logics conceived by the designer for each player and for the room as a whole in an expressive notation (Figure 3).

4.1 Overview

The state engine is based on a continuous dynamical system modeled over a simplicial complex, and coupled to the activity of players in the environment (D1). The simplicial complex represents an N-dimensional *metaphorical space*, the vertices of which correspond to elementary conditions imagined by the designers, such as *Intrude*, *Feed*, or *Reveal* in the TGarden design (see Figure 3 for an example of a sketch of a metaphorical space produced)

The instantaneous state of a player within the system is represented by a point in this metaphorical space, along with a region in a *sensor space* defined by the information sensed from the environment plus any features derived from the sensed data. Each simplex corresponds to a valid combination of elementary conditions; the topology of the simplicial complex defines a narrative landscape, defined by the designers, which conditions the possible evolution of the experience within the performance space.

The trajectory followed by a player's state is a path in the simplicial complex underlying the model. Intuitively, each player's state is treated as a particle with some inertial mass, evolving according to laws of classical mechanics. N-dimensional forces are applied to the player's state, using energy derived both from sensors and current state. The inclusion of energy derived from both the player's activity and the location in the simplicial complex allows the system to evolve continuously according to pre-designed dynamics as well as player movement.

Although the sensor data changes rapidly in response to the user, which may cause immediately perceptible changes in the visual

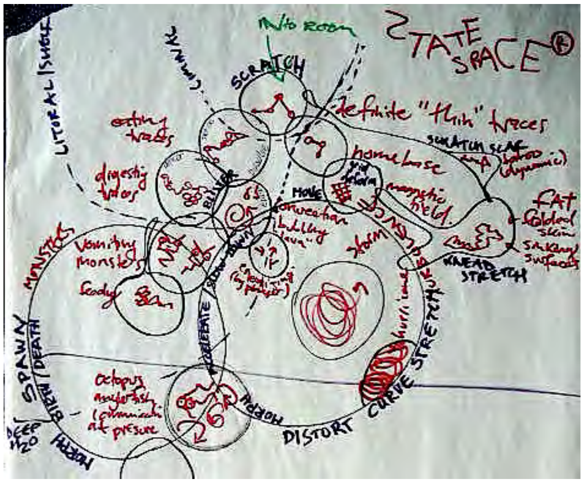

Figure 3: Designers' metaphor space is a simplicial complex.

and auditory landscape, the metaphorical space evolves relatively slowly. When a player state moves from one simplex to another, a perceptible change in the character of the output may also begin to occur as a different set of fundamental states within the metaphorical space become active.

4.2 Player and Room State Spaces

In Ozone, each player is assigned a total state which is inferred from sensor data and metaphorical state. Similarly, the room as a whole is also assigned its own state; formally the room is treated as the zero-th player. The metaphorical state is a point on one factor of the model space of our dynamics system, and it evolves continuously, in a way that is determined by their movements in the space relative to the design decisions about the model structure. By design decisions we mean the choices made by the environment's designers about the implemented state space topology and the parameters governing its behavior.

In this section we describe the player metaphorical state space, the sensor state space, the energetic model which glues the sensor and metaphorical spaces, and the dynamics governing the evolution of player states. These correspond to the data structures, parameter assignment, and time-evolution algorithms underlying our system.

In our applications, we model the dynamics of each player over a base finite dimensional topological vector space space Γ_p of possible metaphorical states, describing the character or color of the environment's response to a player which is in that state. This descriptive state space is one factor of the complete state space M_p for each player, which combines Γ_p with the space S_p of possible sensed activity for that player within the environment, or in other words, the space of sensor data which the system perceives:

$$M_p = \Gamma_p \times S_p \ , \tag{1}$$

which is only to say that the complete information the system maintains about a player p is given at any instant by a position in Γ_p, a set of sensor data, or point in S_p, and associated time-derivatives. The complete states space is the product of those describing each player:

$$M = \prod_{p=0}^{N_p} M_p = \prod_{p=0}^{N_p} \Gamma_p \times S_p. \tag{2}$$

where N_p is the number of players. Our applications have been built for two to six players but it is easy to accommodate any number.

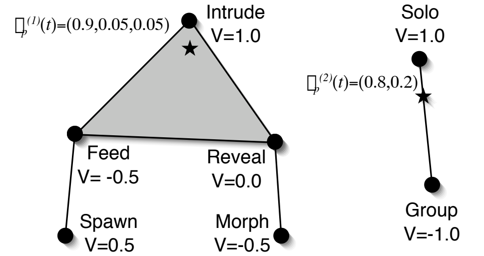

Figure 4: Human player state topology.

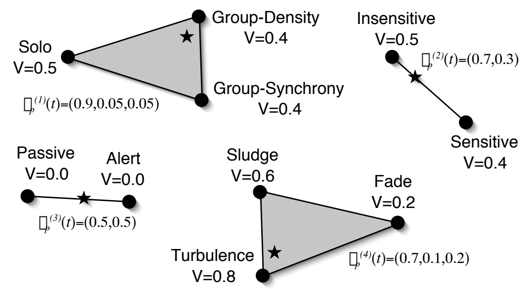

Figure 5: Room state topology (see Figure 3)

This model provides a representation for the dynamics of the media environment that captures the high level semantics of the designers and is at the same time a representation that can be effectively computed. (D4)

We describe both the base and sensor spaces in more detail in the following sections, and subsequently the parametrized energy model through which these two spaces give rise to the system evolution. Figure3 is a sample diagram from the preliminary player state topology written by the designers of the TGarden 2001 system, and Figures 4 and 5 show the machine representation of the same metaphorical design as a simplicial complex of a piecewise-continuous topological space.

4.2.1 Metaphorical State Space

The base space Γ_p defines the possible metaphorical states which the player p in the interactive environment may inhabit (D6). Our current player configuration is shown in Figure 4. p could refer to the room itself; again, we think of the room in this context as another player. Our current room configuration is shown in Figure 5. In our typical configuration, the human player descriptive spaces are all identical, and one may think of the corresponding states as several points on a common human player space. When Γ_p has more than one connected component,

$$\Gamma_p = \Gamma_p^{(1)} \cup \Gamma_p^{(2)} \cup \cdots \tag{3}$$

the player state γ_p is given by a point on each, $\gamma_p = (\gamma_p^{(1)}, \gamma_p^{(2)}, \cdots)$. In the model described below, the sub-states on each connected component behave entirely independently, and may be treated formally as separate players. As a result it suffices to describe the case in which Γ_p has a single connected component, which we will assume in the rest of what follows.

Each space Γ_p of metaphorical states is built up as follows. A collection of N fundamental states $v_k,\ i = 1 \ldots n$ is chosen, and each is named after an elementary condition or scene as imagined by the designers, such as *Intrude*, *Feed*, or *Reveal*. In the player state topology in Figure 4, the pure state named *Reveal* corresponds to the condition of the player appearing to skate across the surface of a magma or ocean, leaving only simple marks on the apparent surface of the projected fluid imagery.

A player state $\gamma_p(t)$ representing a player p at a time t is given by a normalized set of weights for the mixture of states which that player is occupying,

$$\gamma_p(t) = \sum_{j=1}^{N} \lambda_j(t)\, v_j, \tag{4}$$

and the combination of fundamental states which describe it is convex,

$$\sum_{j=1}^{N} \lambda_j(t) \equiv 1\ . \tag{5}$$

The player state $\gamma_p(t)$ determines the metaphorical evolution of that player through the lifetime of that instance of the system. This domain of states is modeled by a set of mixtures, rather than by a discrete graph of nodes and arcs, in order to allow it to interpolate continuously between 2 or more states, corresponding to continuous and rich changes in the environment (D5).

A player only inhabits certain sets of states simultaneously, so Γ_p is restricted to certain permissible n-fold combinations, or *simplices* $\sigma_{jkl..}$, where

$$\sigma^n_{jkl...} = (v_j, v_k, v_l, \ldots) \tag{6}$$

is a simplex spanning the states $v_j, v_k, \ldots v_l$. A player state may occupy a positive mixture

$$\lambda_j(t) > 0, multiindex\, j = k_1, k_2, \ldots k_n \tag{7}$$

only if the corresponding simplex $\sigma_{k_1 k_2 \ldots k_n}$ is contained in Γ_p. At any instant, the player state describes a point on the simplex spanned by the states it is inhabiting. This player descriptive state (or the set of such states) is what is evolved by the dynamics engine, and Γ_p is just the connected union of the simplices of the system, in other words the *polyhedron* of the simplicial complex (D6, D7).

The interaction designers specify the set of fundamental states v_k and metaphorical associations to each, with co-occurrence relationships determining exactly which mixtures of fundamental states, or simplices $\sigma_{ijk..}$, exist in the system (D4). Transitions can be defined at simplex boundaries (D7). This set of boundary conditions between simplices in the complex determines a graph $g(\{\sigma\})$ between the simplices contained in Γ_p. When necessary, we elevate g to a weighted graph so as to allow a mechanism for mediating situations in which more than two simplices share the given boundary, and make it directed in case asymmetric transition relations between neighboring simplices are desired. The weighted, directed graph g is represented by a matrix of floating-point values. Each entry $g(\sigma_I, \sigma_J)$ of this matrix is zero if the boundary from σ_I to σ_J does not exist, or if the transition is forbidden, and greater than zero otherwise.

To summarize, the domain Γ_p of descriptive states $\gamma_p(t)$ for each player p consists of the polyhedron of a simplicial complex built out of vertex representatives of the ontological states of the system, and determining which of those states may be simultaneously active, combined with the set of boundary relations $g(\{\sigma\})$, describing a pseudo-narrative topology determined from design decisions regarding which states meet in a particular composition. A player's trajectory $\gamma_p(t)$ for $t \in [t_1, t_1]$ is a path in Γ_p determined by a physics which we describe in the following sections.

4.2.2 Sensor Space

Associated with each state γ_p is a space S_p of possible sensor features. The sensor data is used to drive the dynamics, to be described below, in response to the contingent activity (D5). The data for each player p consists of a vector of real valued parameters $s_\mu(t)$, obtained from hardware sensors in the room and their derivative features, we use very simple derivative features in our current work, but the model supports any features that can be represented as a time sequence of vectors of floating point values.

The parameters are updated in real time, at a rate which represents the movements at a resolution sufficient for the media synthesis components of the system. This resolution frequently exceeds the requirements of the dynamics model itself, because the effective integration time of the dynamics is significantly longer than than those of the media synthesis components.

When the human narrative spaces also coincide (as has always been the case in our implementations), one can think of these states as particles moving on a common, piecewise-linear domain

$$\Gamma_{\text{common}} \cong \Gamma_1 \cong \Gamma_2 \cong \cdots\ . \tag{8}$$

In cases where the human player states respond to similar sensor data, we have sometimes found it useful to imagine a single sensor space in which the sensor data of each player takes its values. One may then consider the state space as consisting of N_p copies Γ_{common} with a uniform player sensor space S_p.

4.3 Energetic Model

For each player, we engineer an evolving energy landscape over the state topology, and let the player state move as a massive particle on Γ_p, evolving according to the laws of classical mechanics. It is this movement which is recapitulated in the changing character of the output and responsivity of the associated media synthesis instruments.

The potential portion of this energy arises through the coupling of the sensor data acquired from each player to the fundamental states in the system. To a given point $\gamma \equiv \gamma_p(t)$ on the base state space Γ and sensor data vector $s \equiv s_{\alpha,p}(t)$ of S_p is associated an energy given by:

$$U[\gamma] = \sum_{v_k \in \sigma(\gamma)} H_\Gamma(\gamma, v_k)\, H_S(s)\,, \tag{9}$$

a sum over the sites v_k of the current simplex $\sigma(\gamma)$. In this sum, H_Γ gives the energy dependence on the position γ relative to the pure state sites v_k, while H_S is designed to give the energy of the player sensor data vector s relative to data assigned at v_k. A wide variety of energetic models based on (9) are possible. We describe below in more detail the Γ and S dependent contributions which we have found useful. Figure 6 illustrates the model structure and parameters.

4.3.1 Sensor Coupling, H_S

Our sensor energetic coupling has the form:

$$H_S(s) = \exp\left(-\beta E_k[s] + g_s \phi_k\right) + g_V V_k\ . \tag{10}$$

The state-dependent contribution ϕ_k controls the local scale of the sensor contribution to the energy, and the static potential V_k is included to give the media choreography system a background dynamics independent of the sensor activity, in a way that is capable of lending it dynamic tendencies independent of the player sensor data. The coupling constants g_s and g_V control the global relative scale of these contributions and the system's sensitivity to preassigned (H_Γ) or activity-dependent (H_S) energies.

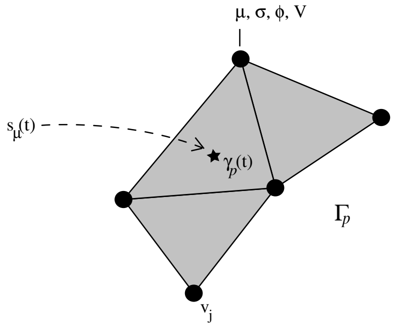

Figure 6: The simplicial model

$\gamma_p(t)$ is the current state. $s_\mu(t)$ is the player sensor data vector. The data associated with each elementary state are the nominal sensor value μ, the variance σ, the static potential V, and the local scale of the sensor contribution to energy ϕ.

Choosing a quadratic dependence

$$E[s] \sim (\mu_{\alpha k} - s_\alpha)^2 \tag{11}$$

because it represents the leading-order of a generic potential with a minimum at the given sensor mean, the inferred player state $\gamma_p(t)$ drifts according to forces arising from the discrepancy between observed sensor features s_α and mean sensor values $\mu_{\alpha k}$ for the fundamental states v_k (D6).

The Γ-dependent contribution to the energy at a player position $\gamma_p(t)$ is given by restricting to the current simplex and choosing any potential function on it. We compute in convex coordinates λ_k on the simplex, and refer the standard embedding of the N-simplex in R^N as the constraint surface

$$\sum_{i=1}^{N} \lambda_k = 1 \ , \quad 0 \leq \lambda_k \leq 1 \ . \tag{12}$$

This implies that the N-dimensional components of any vector fields on the simplex sum to zero, so they remain tangent to it. For example, choosing the inverted quadratic potential

$$H_\Gamma(\gamma, v_k) = -C(\lambda_k^2 - \lambda_k/N) \tag{13}$$

yields the simple linear force $F_k = C(\lambda_k - 1/N)$, where N is the number of states in the simplex.

Dual to the last example, one may consider the quadratic potential

$$H_\Gamma(\gamma, v_k) = C\,(1 - \lambda_k)^2 \ , \tag{14}$$

where $\lambda_k(t)$ are the weights giving the component along the pure state v_k of the player state γ. This is a harmonic potential centered at the pure state site $\lambda_k = 1$, and the force it gives rise to is restorative and proportional to the displacement from v_k. This force does not lie tangent to the simplex, and one must project out the normal portion in R^N.

A class of forces which are manifestly tangent to the simplex are those computed from the Euclidean distance on R^N,

$$d_{\mathrm{EU}}(\gamma, v_k) = \sqrt{\sum_j (e_k - \lambda_j)^2} \tag{15}$$

where e_k is the unit vector which is 1 along the k dimension, and zero elsewhere. The associated distance vector lies on the line between the player state and pure state, and as a result any *central* potential such as

$$H_\Gamma(\gamma, v_k) = C\, d_{\mathrm{EU}}^p \tag{16}$$

gives rise to a force which lies tangent to the simplex. We have for example taken $p = -1$, which gives rise to an inverse-square force law.

The last two potentials give rise to forces having discontinuities across simplex boundaries, and consequently the momentum has discontinuities. But the player state trajectory is continuous. We will describe the dynamics further in the next section (D7).

4.4 Dynamics

In this section, we describe how the model evolves its inferred player states in response to the player sensor data. Evolution proceeds by integrating the first-order equations of motion (20) and (21), including contributions from the gradient of potential and kinetic energy terms.

Using a total potential energy $U[\gamma]$ of the form described in the previous section, the force is computed from the gradient of U,

$$F[\gamma] = -\nabla U - \xi\, \dot{\gamma}_p(t) \ , \tag{17}$$

where we included in addition a damping force depending on a coefficient ξ, and the time derivative $\dot{\gamma}_p(t)$ of the player state, which is useful, for example, for suppressing oscillatory modes of the system.

The force law in convex coordinates is given by

$$F_k = \frac{-dU}{d\lambda_k} - \xi\, \dot{\lambda}_k \tag{18}$$

For example, with an inverted harmonic potential H_Γ, one gets (temporarily setting the background couplings $g_s, g_V = 0$ for simplicity) a force proportional to

$$F_k(t) = (\lambda_k(t) - 1/N)\, e^{-\beta E_k[s_\alpha(t)]} - \xi\, \dot{\lambda}_k(t) \tag{19}$$

with E_k computed as in equation (10).

The dynamics are determined by Newton's second law, $F[\gamma] = m\ddot{\gamma}$. We evolve each player state by means of the first order versions of this, in components:

$$\dot{\lambda}_k \rightarrow \dot{\lambda}_k + \frac{1}{m} F_k(t) dt \tag{20}$$

$$\lambda_k \rightarrow \lambda_k + \dot{\lambda}_k dt \tag{21}$$

where dt is the integration time step, and m is the mass of particle representing the current state.

Framing the dynamics this way gives, in addition to the parameters assigned by the designer at the time of system initialization, several real-time controls over the evolution of the physical model: the mass and damping of each player state, m and ξ, the coupling constants g_s, g_V of equation 10, and the sensor sensitivity β.

The palpable magnitude of certain of these parameters is constrained by the desired timescale for evolution for the autonomous and non-autonomous motion on Γ. The potential energy is the designer-imposed background energy landscape assigned to meaningful states, eliciting evolution even in the absence of sensor data. For example, the typical time-scale governing the motion of a player state of mass m in a harmonic potential of magnitude $g_v V_k$ is

$$T \sim \sqrt{\frac{m}{g_v V_k}} \tag{22}$$

In this case, one chooses m, g_V and each potential V_k such that the autonomous processes of the system have the desired meso/macro time-scale behavior. Parameters governing the motion in response to sensor data must be adjusted separately such that the typical

Figure 7: Ouija Experiments
Dancers under projection of live brushwork by Filip Radonjik, composited with processed video of their own movement.

time-scale for that response has the desired character, and it is easier to do so empirically, when the system is set up. The result of this procedure is to constrain the set of ranges of the dynamical parameters to what is useful for the system's intended response.

5. RESULTS

Ozone presents a fundamental, systematic alternative to scripts, timelines, if-then logic, and FSM's for conceiving and shaping the behavior of a responsive environment, an alternative that robustly accommodates improvisation in rich media environments. Over the past 5 years, we have tested and refined this approach in sizable, concrete artistic applications. (See documentation videos: http://vimeo/tml.)

OUIJA. Ouija (Figure 7) is a phenomenological experiment about gesture, intention, and collective vs. individual agency. The questions we ask are: (1) When is a gesture intentional vs. accidental? How can the system – the people as well as the media – evolve accordingly? (2) When is a gesture individual to a particular body and when is it collective? And how can the system evolve accordingly? Context, history, expectation, and intent all condition these questions, so the answer does not lie solely in the sensor data's instantaneous values. All the agents in an event – the human actors and spectators as well as the computational media – respond to each other concurrently. As with other Ozone applications, Ouija focusses primarily on the first-person experience of the mover, rather than third-person spectator experience.

In 2007, choreographer Michael Montanaro[9] designed a series of Ouija experiments cast in the form of structured improvisation exercises familiar to dancers. Under these conditions, we witnessed the construction of an entirely non-verbal pidgin that seemed to make sense to the participants but not to the observers.

COSMICOMICS. The Cosmicomics installation[10] used continuous state input to direct both sound and video response to pedestrian behavior in a lobby space. Camera-based motion tracking informed the state system to increase or decrease the speed of time progression into successive states. The sound design made use of both the conditioned motion data (smoothed over a 2-3 second time window) as well as the overall state vectors, so sonic responses were both immediate and synchronized with the video projections reflecting the current "mood" of the system. In addition to this responsiveness at multiple timescales, mappings to sound parameters were similarly first- and second-order: state changes modified not only the intensity of various sonic "characters" via fades and effects, but also the degree and method with which the motion data transformed these characters.

[9]Michael Montanaro, Chair, Contemporary Dance, Concordia University. Soo-yeon Cho led these experiments as improvisatory movement exercises, with four dancers and members of the Topological Media Lab in June and July in the Hexagram Blackbox. http://vimeo.com/10828473

[10]Cosmicomics: http://vimeo.com/10825868

REMEDIOS TERRARIUM. In 2008, the Remedios Terrarium[11] installation extended this model to direct the behavior of three separate sound fields in a gallery space, incorporating a method to capture acoustic sound and replay it within the sound system's own logic. While camera and state data controlled multiple synthesis generators using strategies similar to Cosmicomics, the system also detected acoustic events from a high-traffic hallway according to pre-calibrated thresholds. As this hallway was adjacent to a resonating glass-wall, the system captured events only when it was not producing interfering sound in the same space. In other states, these events were transformed and replayed into a sound field adjacent to another high-traffic area, a busy exterior sidewalk. For simplicity, this logic of deciding on events and how/when they should be replayed was kept local to the sound system. However, if this logic was instead implemented by the state system and fed the necessary raw data from the sound system — in this case, event capture timing and raw system output levels for feedback monitoring — it would be possible to coordinate this behavior in tandem with other media systems. There would also be the benefit to the designers of having more potential input data for the state composition.

FRANKENSTEIN'S GHOSTS. The Frankenstein's Ghosts research and production[12] with a professional ensemble of musicians and dancers has shown opportunities for the state-based dynamics. The proscenium-based, well-rehearsed performance is radically different from an installation. However, as the production design calls for transformation of — rehearsed and improvised — musical signals[13] as well as video and lighting controlled by both live and pre-composed means, there is are potentially interesting applications for state-driven behavior. The state system would not be applied the same way in such a scenario where overall composition is still completely predictable, but rather as a flexible mediating layer between media, where it is possible to create rich, changing scene environments whose media have inter-related behaviors according to rehearsed, but complex, dynamics.

6. DISCUSSION

We have described a media choreography system which statistically interprets sensor data generated by body gestures and uses the reduced data to steer a dynamical system to hint the realtime synthesis of video and sound. Using a physical model to generate fine-scale sound and video images, and to track and generate changes of state at scales perceivable and meaningful to human designers and players significantly opens up possibilities for richer responsive environment behaviors than what can be sustained by procedural scripts or aleatoric techniques. This physics-based approach offers an answer to a challenge standing since the 1960s from the experimental performance and entertainment application domain to discover ways to generate aesthetically interesting "script-free" dynamics and interaction patterns without resorting to random processes, since many procedures in fictive and imaginative works, although less than fully specified as deterministic processes, nonetheless do not seem to be random.

Another advantage of our approach is that the dynamics computed on designer-defined states matches concisely and expressively the designers' metaphorical notation. Our model for computing dynamics on continuous state space topologies also provides a uniform way to track user action and respond both discretely and continuously. The model also provides a representation of the state of

[11]Remedios Terrarium: http://topologicalmedialab.net/ remediosterrarium/, http://glia.ca/conu/TML/remediosTerrarium-03-18-08/imageDetrius.swf

[12]Frankensteins Ghosts samples: http://vimeo.com/10835752

[13]Cello WoMAX hybrid improvisation http://vimeo.com/10824436

the media space that can be observed, and interacted with, by the designers, to facilitate both debugging and refining of the design.

Furthermore, by factoring metaphor state computation from the audio and visual media synthesis, the system can provide choreographic hints to the media synthesis instruments and still stay out of the way so they can generate media in response to the sensor data as fast as they can. This parallelism is essential to allow some measure of tangibility and causal coupling to player gesture.

The media choreography system was conceived as a modular framework of independently written media instruments, accommodating a general class of dynamical installation-event structures. Therefore it avoids some legacy problems common to artist-created hardware / software platforms written to support one installation. Over years we have smoothly migrated across four generations of physical movement tracking hardware and entirely new media programming environments (NATO to Jitter, SuperCollider to MSP), with applications over a large range of installation-events, from responsive environment to gallery installation to dance performance.

Based on our experiments with dancers and musicians, we are extending our work to architectural applications (Figure 8) in the built environment, and to ubiquitous, mobile applications.

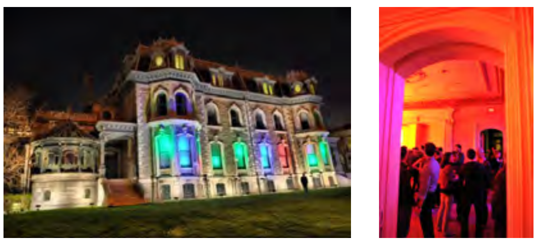

Figure 8: Active lighting by M. Sutherland, N. Navab, & TML *(A) exterior LED array, (B) induced interior field; Canadian Centre for Architecture 2009. Photo: Benoit Desjardins.*

7. ACKNOWLEDGEMENTS

The 2007-2010 Ozone System was implemented at Concordia's Topological Media Lab by Michael Fortin (graphics and computational physics), Jean-Sébastien Rousseau (realtime video), Timothy Sutton, Navid Navab (realtime sound), Morgan Sutherland (state engine composition and camera processing), Harry Smoak (lighting, set, and experiment design), and Emmanuel Thivierge (state engine abstractions). This work was supported by the Canada Fund for Innovation, Canada Research Chairs, Social Sciences and Humanities Research Council of Canada, Hexagram, Concordia University: Faculty of Fine Arts, the Office of the Vice President of Research and Graduate Studies, and Office of Research.

We thank the artists and engineers who worked on the creative applications that have motivated this system over the years. Yon Visell wrote the original implementation of the state engine. The TGarden team presenting TG2001 at Ars Electronica and V2 symposia included Sha, Maja Kuzmanovic, Chris Salter, Laura Farabough, Joel Ryan, Nik Gaffney, Evelina Kusaite, Yifan Shi, Cocky Eek, Yon Visell and members of the sponge and FOAM networks. These productions were supported by the GVU and New Media Centers at Georgia Tech, Starlab, Banff New Media Institute, V2, Ars Electronica, the Langlois Foundation, and other cultural foundations. We thank Georgia Tech's College of Computing for supporting the development of the tgvu media choreography testbed in 2001-2003, by Yoichiro Serita, Vincent Fiano, Erik Conrad, Jehan Moghazy, Jonathan Shaw, David Demumbrum of the TML.

8. REFERENCES

[1] Aaron Bobick, S. Intille, J. Davis, F. Baird, Claudio Pinhanez, L. Campbell, Y. Ivanov, A. Sch*ö*tte, and A. Wilson. The kidsroom: A perceptually-based interactive and immersive story environment. *Presence 8: Teleoperators and Virtual Environments*, pages 367–391, November 1996.

[2] I. I. Bukvic, J. Kim, D. Trueman, and T. Grill. munger1 : Towards a cross-platform swiss-army knife of real-time granular synthesis. In *ICMC Proceedings*, 2007.

[3] Lucy Bullivant. *Responsive Environments: Architecture, Art and Design*. Victoria and Albert Museum, 2006.

[4] Center for New Music and Audio Technologies. Open sound control. Technical report, CNMAT, University of California, Berkeley, 2002.

[5] M. Garelli and M. Gribaudo. Performance Analysis of the ARIA Adaptive Media Processing Workflows using Colored Petri Nets. *Electronic Notes in Theoretical Computer Science (ENTCS)*, 232:55–73, 2009.

[6] Neil A. Gershenfeld. *When Things Start To Think*. Henry Hold and Company Inc., 1999.

[7] Usman Haque. The architectural relevance of gordon pask. In Lucy Bullivant, editor, *4dsocial: Interactive Design Environments*, pages 54–61. London: John Wiley and Sons Ltd., 2007.

[8] S. E. Hudson and G. L. Newell. Probabilistic state machines: dialog management for inputs with uncertainty. In *Proceedings of UIST 1992*, pages 199–208, 1992.

[9] L. Kovar and M. Gleicher. Simplicial families of drawings. In *Proceedings of UIST 2001*, pages 163–172, 2001.

[10] Myron W. Krueger, T. Gionffrido, and K. Hinrichsen. Videoplace, an artificial reality. In *Proc. CHI85 Human Factors on Computing Systems*, pages 35–40. ACM Press, 1985.

[11] Tom Ngo, Doug Cutrell, Jenny Dana, Bruce Donald, Lorie Loeb, and Shunhui Zhu. Accessible animation and customizable graphics via simplicial configuration modeling. In *Proceedings ACM SIGGRAPH Conference*, pages 403–410. ACM Press, 2000.

[12] M. Ouimet and K. Lundqvist. Timed abstract state machines: An executable specification language for reactive real-time systems. Technical report, May 2006. Technical Report ESL-TIK-000193, Embedded Systems Laboratory, Massachusetts Institute of Technology.

[13] S. G. Parker, D. M. Weinstein, and C. R. Johnson. The scirun computational steering software system. *Modern Software Tools in Scientific Computing*, pages 1–40, 1997.

[14] L. Peng, G. Kwon, Y. Chen, K.S. Candan, H. Sundaram, K. Chatha, and M.L. Sapino. Modular Design of Media Retrieval Workflows Using ARIA. *Lecture Notes in Computer Science*, 4071:491, 2006.

[15] Lina Peng, Gisik Kwon, K. Selçuk Candan, Kyung Ryu, Karam Chatha, Hari Sundaram, and Yinpeng Chen. Media processing workflow design and execution with aria. In *MULTIMEDIA '05: Proceedings of the 13th annual ACM international conference on Multimedia*, pages 800–801, New York, NY, USA, 2005. ACM.

[16] Claudio Pinhanez and Aaron Bobick. It/i: A theater play featuring an autonomous computer graphics character. In *Proceedings of the ACM Multimedia 98 Workshop on Technologies for Interactive Movies*, pages 22–29. ACM Press, 1998.

[17] K. Ryall, J. Marks, and S. Shieber. án interactive constraint-based system for drawing graphs. In *Proceedings of UIST 1997*, pages 97–104. ACM Press, 1997.

[18] Xin Wei Sha, Michael Fortin, and Jean-Sébastien Rousseau. Calligraphic video: a phenomenological approach to dense visual interaction. In *MM '09: Proceedings of the seventeen ACM international conference on Multimedia*, MM '09: Proceedings of the seventeen ACM international conference on Multimedia, pages 1091–1100, New York, NY, USA, 2009. ACM, ACM.

[19] Flavia Sparacino, Glorianna Davenport, and Alexander Pentland. Media in performance: interactive spaces for dance, theater, circus, and museum exhibits. *IBM Systems Journal, Vol. 39, Nos. 3 and 4*, 2000.

[20] R. van Liere, J. D. Mulder, and J. J. van Wijk. Computational steering. In *HPCN96*, pages 35–40, 1996.

[21] Sha Xin Wei. The TGarden Performance Research Project. *Modern Drama*, 48:585–608, 2005.

[22] Sha Xin Wei. Poetics of performative space. *AI and Society*, April, 2007.

[23] Sha Xin Wei, Chris Salter, and Laura Farabough. Sponge: Unstable surfaces. *ec/artS, Issue 2, 2000*, 2000.

[24] Sha Xin Wei, Tim Sutton, Jean-Sébastien Rousseau, Emmanuel Thivierge, Olfa Driss, Josée-Anne Drolet, and Michael Fortin. Cosmicomics responsive video installation 2007.

[25] Sha Xin Wei, Tim Sutton, Jean-Sébastien Rousseau, Emmanuel Thivierge, Josée-Anne Drolet, and Michael Fortin. Meteor shower responsive video installation 2006.

[26] Jan C. Schacher, Philippe Kocher. Ambisonics Spatialization Tools for Max/MSP. in Proc. ICMC, New Orleans, 2006

[27] Jot, J-M. and Warusfel, O. A real-time spatial sound processor for music and virtual reality applications. in Proc. ICMC, San Francisco, pages 294-95, 1995