7 Practices: Apparatus and Atelier

This chapter is concerned with the “hows”—some of the practical and instrumental matters involved in the kind of artistic or critical work informing or inspired by the previous chapters. How would you build an apparatus that could even begin to respect a nonanthropocentric conception of living matter, and yet condition a built environment that could sustain ethico-aesthetic experiment? How would you do that at the level of engineering? What sort of organization in what sort of institutional ecology, and what ecology of practices, would be adequate?

We start with a specific technical application of some of the mathematics, as a way to build an apparatus for playful improvisation of individually and collectively charged movement and gesture. This is the sensing and media synthesis system built with elements of the responsive media installation-event examples populating chapter 3 and other parts of this book. Then we enlarge the scope to the institutional organization and ecology of practices that can house the making and evolution of such experimental research and creation.

Ozone, a Media Choreography System

In this section, I give a technical description of a responsive media software framework that is informed by (1) observing how composers compose a potential event by sketching metaphorically, and (2) the continuous classical mechanics of fields and motion actualizing potential fields of metaphor.

This “media choreography” system, called Ozone, is designed to support dense, continuous play. Each of these terms has a sense taken up in the earlier chapters of this book. By play I mean unschematized, improvisatory but not random activity (from the point of view of the performer). These qualities can be applied not only to the millisecond scale of producing real-time media in concert with inhabitants’ action, but also to technical ensembles for gesture (cf. Doug Van Nort’s work), composition of event (Ozone), and perhaps coding, and to a social nexus in a field of institutions, such as the Topological Media Lab.

Figure 7.1
Life in the Topological Media Lab atelier-laboratory.

When we creators gathered to compose the “look and feel” for the Frankenstein’s Ghosts experimental performance or the TGarden play space installation-event, it was not at all clear how we should go about shaping the environment, especially in a collective circle of peers. We set aside structuring schema such as scripts, scores, and soundtracks. In general, we cannot assume that the visitors can play or reperform any musical patterns with any skill. Using a multilinear track-based logic puts a cage on potential action and leaves little room for improvised expression and surprise. Having designers make up stories about what a potential visitor would experience in a responsive environment is not so helpful beyond a certain point, because every single visitor’s narrative, however skillfully imagined, is just a particular one-dimensional trajectory through an infinite-dimensional space of intensities and fields. No finite set of these narrated trajectories could ever add up to a thick description of the environment that could guide its design, much less its construction. But a fictive physics articulated in the spirit of equations of state could suggest potential dynamics as microcosmologies. We can breathe life into an imagined cosmology by describing not its exact progression but its tendencies, a potential field of possible states through which an event could evolve relative to a set of affective or metaphorical states. Composers can sketch in their own terms how the event should feel upon entry or exit; some moments of intensity or repose, or mystery, or ambivalence and multiplicity; what states could overlap or blend with what other states; what are one state’s tendencies toward evolving into other states; what sort of physically observable conditions or activities are associated with a given state; and so forth. This way, designers can fluidly imagine and revise how the responsive environment could evolve not just from the point of view of a particular inhabitant’s trajectory, but for any and all of the inhabitants in any condition (see figure 7.2).

How can we marshal all the *concurrently running* and *dense* media processes (not prefabricated sound or visual objects!) in concert with both prior aesthetic intent and contingent ethico-aesthetic action? Since the media textures may be very finely crafted according to musical and visual design, the system should support the control of richly structured transformations rather than just “random” sequences of cued media. At the

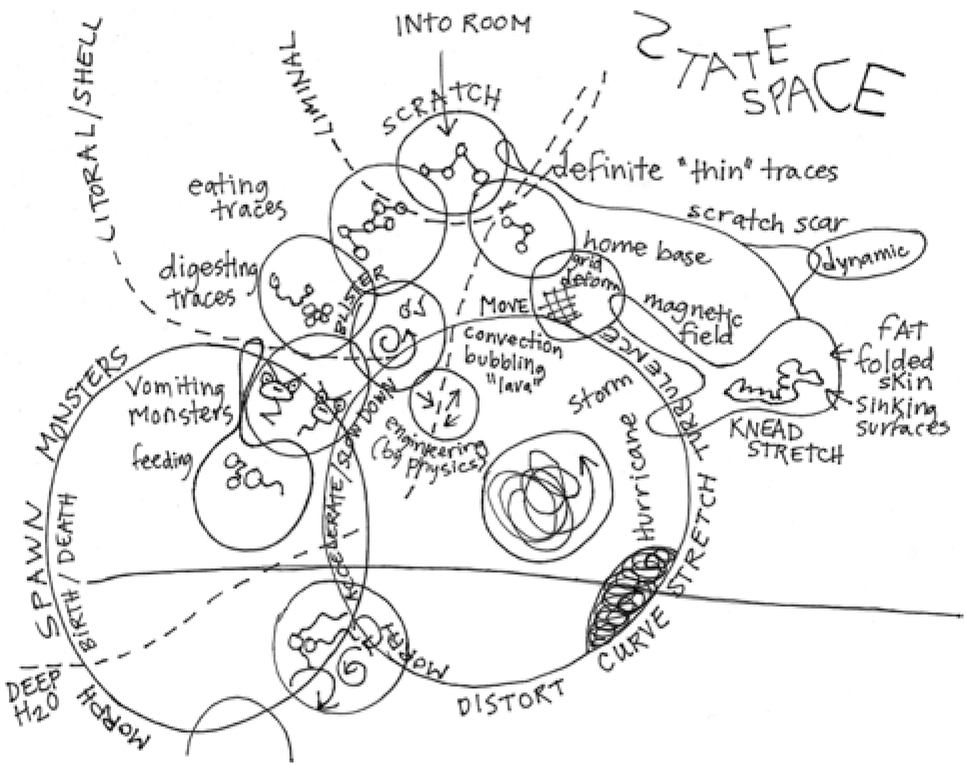

Figure 7.2
Designers' metaphorical sketch of a field of potential events, produced during a design session at Banff New Media Institute, 2000. Image courtesy of Sponge.

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 are too sparse or too complicated to support the dynamical evolution of fully dense media. Representations in conventional procedural languages are too low-level to efficiently capture the rich, multivalent or inconsistent semantics of the artists' metaphorical designs. We can pitch the system'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.

A *media choreography system* is a set of software (and hardware) frameworks that extracts features from sensor data tracing what is going on in a physical space, and creates or modulates ambient media (video, sound, lighting, kinetic material, or objects) in real time, concurrently with that activity. The mapping from activity to media dynamics can follow one of many strategies: (1) fixed timeline, such as cue-based systems, which impose a fixed, predesigned sequential order on events that can only be repeated; (2) if-then logic, which tends to be brittle and hard to revise on the fly; (3) stochastic methods, which provide a degree of unpredictable variation, but with no ethico-aesthetic *sense* to the distribution—or rather a very particular one: that of the *random*; (4) written scripts, which permit a great deal of interpretation by persons inhabiting the event, but not by the costructuring software or material. It should accommodate all the contingent activities and conditions in the environment, whether due to people, the media, or the environment, as well as any prior intent on the part of the composers or designers of the environment.

I describe a particular media choreography software system called Ozone, implemented by the Topological Media Lab via five generations of development from the prototype TGarden system to the present. Ozone has the key feature that the pseudo-physics and high-level evolution of the climate of the environment can be molded by means of a representation that is legible to dramaturgical and aesthetic composers who do not think like programmers. Given that the experience of a responsive environment evolves qualitatively like a dynamical system, a composer's expressive design for such a responsive space could amount to specifying the environment's pseudo-physics. This way of thinking is quite different from brittle procedural, Boolean, or data-based programming logic.

A subtle difference between an information-theoretic approach to scripting the behavior of a system and the Ozone media choreography system's quasi-physical approach is that the latter bets on a radically modest approach to computational media as dumb matter. By dumb I mean (1) free of language, even the formal procedural programming languages that are operationalizations of the logic that I relinquished early in our experimental work; (2) free of intelligence, as in the cognitivist approaches of symbolic artificial intelligence; and also (3) free of representations of abstract structures like hidden Markov statistical models or 3D polyhedral geometry.

A key common feature of the media choreography of this family of play spaces is that the creators specify not a fixed, discrete set or sequence of media triggered by discrete visitor/player actions, but rather a potential range—a field—of possible responses to continuous ranges of player actions. But in this family, behavioral tempers (or, to use less animistic terms, climates of response) evolve over macroscopic periods of time (minutes), according to the history of continuous player activity.

The multivalence can be articulated as multiplicity of state. For example, a *player* could be described as a soloist or as part of a group. But rather than simply flip between two discrete player states like Solo and Group, it makes expressive sense to generalize this aspect of a *player* to a *continuous* range of "groupness" between extremal Solo and Group states. One advantage of our physics-based approach is that it provides a level of abstraction between the sensors and the rest of the system. The composers can get on with the shaping of experience without freezing on a particular set of physical conditions in some brittle, intricate code that would have to be ripped out and rewritten should they change their minds. For example, once groupness is defined as a concept based on the two metaphorical components Solo and Group, the composers can begin using it in the media synthesis engine even as the detailed choice of sensing modality and feature extraction method mutates in the course of the engineering design and development. 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. *How groupness is related to physical action or configuration can be decoupled from the software or hardware implementation, and varied in the rehearsal process*. The final choice can be made as the compositional and rehearsal processes interweave with the engineering production of a responsive environment.

Assuming a rich, dense responsive environment full of responsive calligraphic video, responsive sound, and active electronic, sounding, kinetic textiles, all responding *concurrently* to activity and conditions in real time, how can a composer hope to shape such a complex environment without becoming a bureaucrat of media, or a tyrant of experience? How can we marshal dozens and hundreds of concurrent media processes creatively? When do some media processes start, and how do they vary in response to other processes in an event? Based on physical dynamics, the Ozone media choreography system is designed and built to support interactive spaces that require the real-time synthesis and coordination of arbitrary streams of video and audio in response to actions by one or more people.[1]

Artistic concerns—which include ethical as well as aesthetic concerns—motivate the following desiderata:

(1) The composer, actor, spectator may be the same body, implying that we focus on first-person experience;
(2) The primary modes of interaction are not based on (isomorphs of) linguistic patterns, but on continuous fields of matter and media;
(3) The participants are always in a common physical place, setting a very high demand for sensuous density and effectively zero latency;
(4) The composer composes not specific event sequences but *meta*-events, or substrates of fields of potential events;
(5) Design for *continuous* experience with the density of everyday settings.

The second and last of these desiderata again motivate a topological versus discrete approach to time-based media.

Technical Features of the Media Choreography System

Given those artistic concerns, I designed the media choreography system to:

(1) Directly accommodate high-level composer semantics for interpreting or responding to player activity;
(2) Incorporate arbitrary continuous as well as discrete evolution of state;
(3) Support low latency responsiveness to sensor data;
(4) Support synthesis of structured and perceptually dense, continuous video and audio with plausible response in real time;
(5) Leverage the intuitive concepts of energy-based physics of material dynamical systems to allow composers to create potential event-landscapes—called *state topologies*—in a way that is idiomatically concise and expressive.

In brief, the media choreography system allows the media synthesis system to use any combination of sensor and metaphorical data to generate meaningful, even compelling, aural and visual patterns on the fly from live or prepared audio or video textures. By integrating the basic metaphoric components of the composers' language with sensor data and derived sensor features in a continuous, real-time evolutionary system, this media choreography framework *simultaneously* lets the composers think in terms of evolving metaphorical states while at the same time constituting a potential-dynamic system that can be evolved by the computer using computational physics and topological dynamics.

The prototype media choreography system I designed was tested under harsh performance conditions in 2001 at the Ars Electronica and V2 by the TGarden consortium, which provided valuable feedback for the design of the system. From 2001 to 2010, nine distinct responsive environments and installations were built using elements of the media choreography system.

The central technical challenge is how to make a navigable and playable responsive media space that has no preassigned interface objects nor prespecified gesture/action sequences. A typical "interactive" installation has a behavior that, however rich in its basic dynamics, basically does not meaningfully change its type: an eternal thunderstorm of particles, for example. But how could we compose a responsive media environment whose behavior qualitatively changes in a meaningful, palpable way according both to the composer's design and to contingent activity?

The *media synthesis instruments* project video and sound into a room as ambient fields, continuously changing according to autonomous dynamics and in response to player activity. As the players interact with the projected sound and imagery over time, they should be able to *invent* gestures that meaningfully shape and control the media. This implies that we avoid prefabricated user interface objects or prespecified gestures but instead allow the construction of manipulable objects out of the media textures themselves.

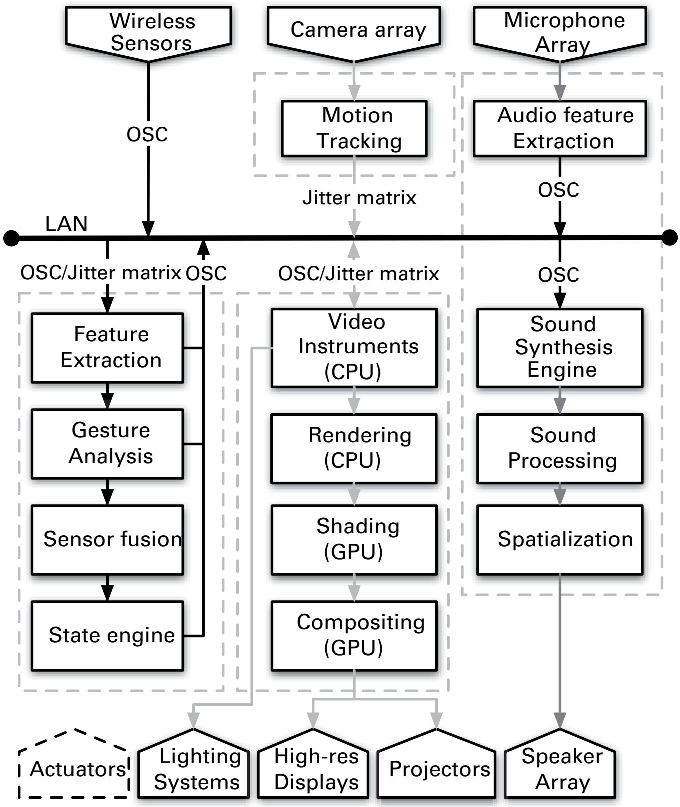

Figure 7.3
Ozone architecture: the architecture of the current media choreography system. Diagram by Morgan Sutherland and Sha Xin Wei.

Hardware and Software Components

Our real-time system is written as a set of components, each on a separate host: (1) a data-routing, statistics, gesture-tracking and media choreography engine written in MAX and C,[2] (2) a sound synthesis engine written in Max/MSP and SuperCollider with externals such as a granular synthesis and DSP analysis, and (3) a video effects engine written in Max/Jitter and C.[3] We chose to implement our real-time media synthesis instruments using Max/MSP/Jitter and SuperCollider because these are among the most expressive, popular, and professionally maintained systems currently used as a lingua franca by media artists and musicians. This gives us the possibility of efficiently incorporating rich musical and visual imagery and dynamics without writing special-purpose effects from scratch for restricted laboratory demos.[4]

Wireless Sensor Networks

In addition to these core system components, we used auxiliary devices to sense activity. For example, to track players' locations in the 2001 version of TGarden, we used a video camera to follow infrared LEDs worn on the players' heads. This turned out to be more robust than visible-light vision methods, since the sole illumination came from projected video under which players are often camouflaged. We also outfitted each player with two sets of Analog Devices ADXL202e accelerometers, whose output was mapped to a Compaq iPAQ running LINUX and beamed via 802.11b wireless Ethernet to a fixed computer. In later applications, we used the low-overhead TinyOS wireless sensor platforms,[5] 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. The locus of computation can migrate between body-borne or fixed hosts as the tradeoffs between platforms, bandwidths, and power supply evolve. Currently, where possible we use infrared camera-based computer vision to track movement and form to avoid "instrumenting"—attaching a device to—the participants.

All software components communicate using Berkeley CNMAT's OpenSoundControl (OSC) protocol.[6]

Clothing as Interface

Each player is assigned an instrument model which is partially parametrized by their continuous movement. The room/instrument itself evolves over time as it adapts to the players' more or less expert gesture.

There are several reasons why treating clothing as interface is interesting and fruitful. We can design the garments to test ballistics that augment or constrain player gestures and poses in ways that can be mapped to vocabularies of motion, much as a musical instrument's physical features provide idiomatic kinesics that make it learnable. Virtual instrument design[7] and wearability studies[8] indicate that carefully designed physical constraints are crucial to the playability and learnability of a gestural instrument. The composers created the clothing with this in mind.

We started by looking to clothing as the most natural body-borne interface between a human and his or her environment, and working with fabric and garment composers who are willing to experimentally extend or constrain the body to provide a variety of gestural affordances. It is essential for theatrical and psychological purposes that the costumes' fantastical design inspire the players to move freely and feel encouraged not to simply habitually repeat but to improvise gestures. It is telling that when professional dancers visited the TGarden, after a brief period of habituation they moved far more surely than amateurs with the moving video projection and music, but they

Figure 7.4
Gestural instruments based on TinyOS wireless sensor platforms, Ubicomp, 2003. Picture by the author.

tended to do their own disciplined movement rather than listen and coconstruct their motions in concert with the room's response.

Allowing arbitrary improvised gesture implies that the body-worn devices should not impede the players' movement. Therefore, we do not equip the players with head-mounted devices (HMDs, goggles), cameras, or headphones. We chose to use accelerometers[9] with 2 *g* range because they detect the right range of forces of our gestures (from free-swinging limbs to sharp chops of the hand), are small enough to be worn unobtrusively as watch-sized pouches on the body, and could withstand the shock of dance movement. We use wireless broadcast of sensor data because tethers would unacceptably constrain the players' motion.

Feature Extraction

Given a set of sensors, one of the key problems to be solved by any interactive media system is how to reduce the multiple time series of sensor data to extract useful features. The Ozone system performs relatively simple statistical reductions: time-based or sample-based moving-window averages or simple $W^{p,q}$ (Sobolev-type) 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 try to perform the sensor data reduction by an automatic supervised or unsupervised procedure, adapting, for example, a mixture model, or performing independent component analysis (ICA) 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. I believe that a good way to achieve this is to allow the composers to manually specify the sensor feature mapping.

Video Analysis

Most of our video analysis is done live using standard computer vision tools such as optical flow measures or morphological filters from a standard computer vision library via Jitter, but we have built some moving-window background thresholding and adaptive motion extraction extensions.

Audio Signal Analysis

Audio signal analysis[10] modules combine common signal analysis algorithms such as spectral centroid, flux, kurtosis, envelope following, bark amplitudes, mel-frequency cepstral coefficients, and derivatives of these features that can be used compositionally, as receptors from other components of the system, and as emitters injecting biased dynamics into other media elements.

Signal analysis procedures using a priori models can behave unexpectedly due to unpredictable activity in our improvisatory and experimental environment. For example, the lighting or acoustical characteristics of the space could change radically, various stages in the system generate noise, and the media systems themselves generate feedback, both at the level of reproduction (projections and microphone pickups) and of analysis. The behavior of the algorithms implemented in a given configuration can fluctuate considerably by design, complementing the system's overall field state-based behavior.

As an example of a derivative audio feature: a simple "activity" measure calculates the amount of fluctuation of the spectral centroid of a boundary microphone mounted on a conference table, by basic prefiltering and onset smoothing. The fluctuation drives a two-stage cascaded leaky integrator, which advances the system with two "rates" of activity. The dynamical result of the fluctuation flowing through the two integrators is usefully richer than one-to-one linear functions of activity.

For another application, the Remedios Terrarium responsive environment,[11] we built a simple "dreamer" that functioned on minute (or zero) amounts of audio, "listening" for events, capturing these into buffers, and emitting them audibly at a later time, via distorted or otherwise transformed playback. The occurrence and quality of these varied depending on the diurnal state of the system. Storing regular intervals of fast Fourier transform (FFT) snapshots, we can generate continuous "fabrics" or "thumbprints" of the acoustic timbre of a space, and these can be presented in a form derived from the audio and/or can inform other media.

In recent years, we have synthesized acoustic and gesture analysis techniques from IRCAM into our system, notably a grain-based coding of sound corpora by perceptual descriptors (CATART), hidden Markov models revised for sparse data with no training (OMAX), and continuous gesture following (*gf*).[12] This has greatly extended the subtlety of tracking not only acoustic information from ordinary air microphones, but also camera-based optical information as well. Most promisingly, we can now break a bottleneck faced by camera-based methods restricted to the low frame rates of conventional video cameras and video acquisition. Where video acquisition rates were at best on the order of 15 to 24 frames per second, digital audio frame rate is typically 44,000 frames per second, or more. Compared as sensors, therefore, a microphone's temporal resolution is typically 1,000 times greater than a camera's.

Time Scales

While the media choreography engine runs in real time, the architecture of the system effectively operates at three different time scales. In practice, we found three scales of temporal dynamics to be meaningful to the composers: *micro scale* of $O(1/10^2)$-second sensor data (e.g., denoising at the sensor PIC), *meso scale* of $O(1)$–$O(10)$ seconds: composer-specified gestural grain (e.g., the rate at which perceptible changes in the value of a fundamental state like Solo occur), and *macro scale* of $O(10^2)$ seconds: a "narrative" state, the rate at which the system switches between simplices, which people can perceive as the unfolding of an event.

Media Synthesis

I have emphasized the core "state engine" of the media choreography system because of its conceptual implications. But the real-time sound and image synthesis instruments written in Max/MSP/Jitter and similar real-time media programming environments are what create the perceptual richness of the performative event. These instruments are designed to be controlled by multiple data streams from the Ozone network such as the "raw" and "cooked" (numerically reprocessed) sensor data as well as the overall state topology vectors.

Audio Synthesis

The sound instruments include several self-contained designs such as a wind generator and the polyphonic vocal synthesis engine incorporated into the Meteor Shower

(2006) and Cosmicomics (2007) installation-events.[13] But development is primarily focused toward a concatenative synthesis framework via Max/MSP. We prefer concatenative synthesis because of its wide range of timbral possibilities, both in terms of seed sound material (any audio file can serve) and multiple levels of parameter modulation. In fact, every unique configuration of parameters (i.e., "preset") of a module of MSP abstractions constitutes a software instrument with its characteristic sensitivities, temporal behavior, sonic quality, and transform or synthesis logic. We design our parameters to be metaphorically meaningful to the designers (i.e., "sludgy" rather than "xyzz1"). This allows the sound programmers to rapidly create very different sonic behaviors that can be auditioned by a composer. The behaviors can be copied, modified, and extended by a composer more familiar with sound design and Max/MSP.

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 raises the question of which components in the system should be calibrated to which ranges. A granular instrument's metaparameters are designed to operate within a nominal range of the continuous interval [0, 1], but the instrument does not constrain input to this range, always allowing for the possibility that a system will attempt to drive this instrument outside of the expected boundaries.

This decision comes from early experiments involving mappings of various (and variously conditioned) channels of sensor and video analysis input to about a dozen 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.

The use of floating-point data throughout all stages (immediately following data acquisition) of such data "munging" prevents stages that may be very low in dynamic range from discarding potentially useful information through truncation and round-off error.[14]

Video Synthesis

My general strategy is to treat video not as image (a picture of something) but as structured light. And if the video is synthesized by our real-time responsive software, then this structured light behaves in concert with the action or condition of that with which it interferes. Calligraphic video as palpable light field becomes an alchemical substance, a shadow of Heraclitean fire. In the lab, we are trying to create real-time responsive video as a substance that, coupled with computer vision or other more tangible sensing techniques, can be manipulated as a painterly medium.

Rather than rely on primitive oculocentrism, we tap the user's large pool of corporeal intuitions about the behavior of continuous physical material to build interactions with dense visual textures in novel and complementary applications. Every person acquires corporeal intuitions from infancy on, so it seems reasonable to leverage that sort of lifelong, preverbal capacity.

We are building these calligraphic video interfaces as platforms for research in gesture along phenomenological lines. Gendlin complements the logical structure of cognition with felt meaning, which has a precise structure: "Experiencing is 'nonnumerical' and 'multischematic' but never just anything you please. On the contrary, it is a more precise order not limited to one set of patterns and units."[15]

Moreover, categories may be logically but not experientially prior to instances. This is a strong motivation for seeking a non-object-oriented approach to manipulable, active graphics. Human experience is material and corporeal, and is intrinsically structured as temporal processes. (This motivates our turn to dynamical fluids.) Based on fundamental work with immune and nervous systems, Maturana and Varela moved from the discussion of cellular organisms to autopoietic systems, loosely and briefly defined as continuously self-reproducing sets of processes in an ambient environment, whose relationships remain dynamically intact across changes of constitutive matter.

Given that, at the everyday scale, experience is continuously composed of temporally evolving matter, we wish to have an experimental platform for creating objects of experience that do not have to be selected from a preexisting category. For example, graphic objects in our manipulable system must not appear to the user as built out of a preexisting set of geometric primitives. It is essential, of course, that these be manipulable in some improvised way, and essential that these manipulations be continuous in time, to permit us to study the evolution of material form—ontogenesis, to use René Thom's term.[16] We build calligraphic video: video texture that responds to manipulation by human gesture as interpreted from camera-based input—as apparatuses in which we can conduct studies of how humans imagine, create, and perceive dynamical "objects" from fields that are effectively continuous in time and space. Working with continuous fields of video permits us to construct experiments in which objects can be formed by improvised manipulation and allowed to return to general substrates. The manipulations must be as free as possible of class-based tools or menu structures (else they would imply preexisting logical, functional, or geometric categories). The video texture substrate may not appear uniform at all, but it is continuous in space and time. Rather than use arbitrary dynamical systems to animate the responsive video, we choose to study the structure of corporeal-kinesthetic-visual intuition via improvised manipulations of media that leverage corporeal-kinesthetic-visual experience of continuous matter commonly encountered from childhood.

Practically, our strategy is to treat video as initial data for physical models of material like water (Laplace wave partial differential equation [PDE]), smoke (Navier-Stokes PDE), etc. We borrow the best techniques from computational physics that now can be executed in real time. Indeed, since video is typically projected onto two dimensions anyway, we may as well forgo 3D graphics and free up computational resources to compute and present much richer 2D textures.

Over eight years, we have created several generations of real-time video-processing abstractions in NATO, and now in Jitter. At a time when video was mostly edited offline and the tools where designed for such paradigms of use, these video instruments were designed to respond to live video streams with negligible latency because they were intended for live performance.

Using the Max/Jitter framework, we can implement processing instruments on multidimensional dense lattices at video rates, streamed between local networks of hosts to parallelize the computation. An example of a visual instrument is the Meteor Shower (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 independent of the number of attractor points. Over the past two and a half years, we have built an extensive library of CPU- and GPU / Open-GL-based real-time video instruments. We describe the computational physics and strategies for parallelizing on different hardware in a technical paper.[17]

State Engine Continuous State Dynamics

In this section, we present a formal description of the dynamical system which models the meaningful configurations of people and activity understood by our system. This model is intended to compactly capture the metaphorical ontology and dynamical response logics conceived by the composer for each player and for the room as a whole.

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. The simplicial complex represents an *N*-dimensional *metaphorical space*, the vertices of which correspond to elementary conditions imagined by the composers, such as Intrude, Feed, or Reveal in the TGarden's state topology.

To be clear, the Ozone state engine is quite different from the finite state machine (FSM) of classical computer science, based on abstract discrete states. It is modeled on the *continuous* model of potential fields and material dynamics of physics. In the Ozone model, states can overlap; the environment can be in more than one basis state simultaneously. (Here I use "basis" in the sense of a topological vector space and quantum mechanics.) And equally importantly, the state of the environment can jump between discrete states but can also vary continuously "over time."

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 composers, which conditions the possible evolution of the experience within the performance space. (For example, a player's evolution through a connected set of simplices might describe a timeline in which a player passes from some marginally active set of states, through some transitional region, and into one meant to give rise to a more turbulent responsiveness.)

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 predesigned dynamics as well as player movement.

Temporal Dynamics

While the engine runs in real time, the design gives the appearance that it is running at multiple time scales. This allows the system to simultaneously sustain a sense of tangibility based on fine-grain temporal response (e.g., for features derived from sensor data), as well as of global evolution of state (e.g., for the metaphorical state). For example, the sensor data changes rapidly in response to the user, which may cause immediately perceptible changes in the visual and auditory landscape. However, the values of active fundamental states within the metaphorical space change relatively slowly in response to the sensor data and the physical simulation, and thus cause perceptible changes at a slower rate than the sensor data. Finally, 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; however, movement between simplices happens infrequently compared to movement within a simplex (see figure 7.5).

Player and Room State Spaces

In our system, each player is assigned a total state which is a combination of the player sensor data and metaphorical state. Similarly, the room as a whole is assigned a total state; formally the room is treated as the zeroth 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 the players' movements in the space relative to the design decisions about the model structure. By design decisions we mean the

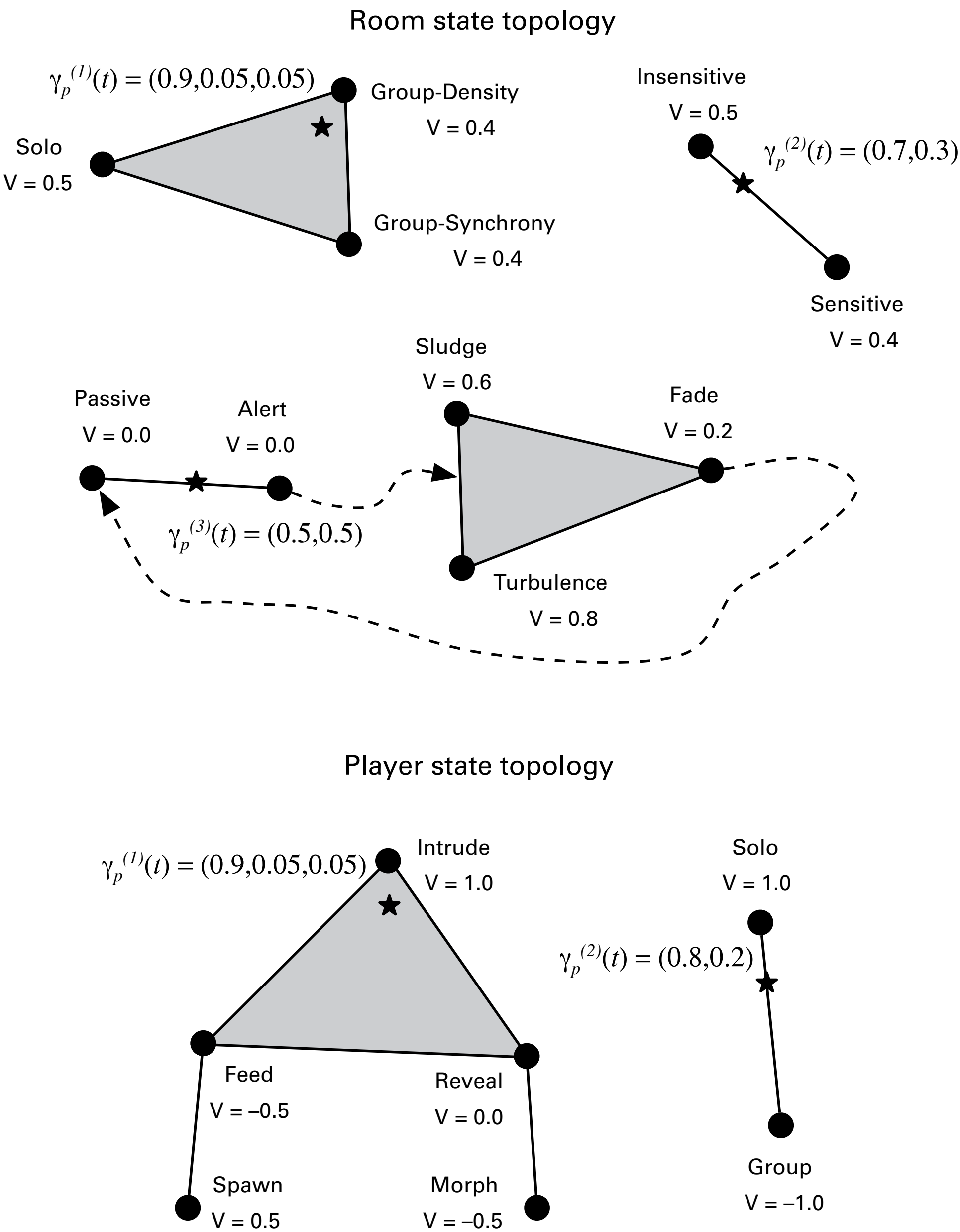

Figure 7.5
Room state topology and person state topology. Diagram courtesy of Maja Kuzmanovic and Sha Xin Wei.

choices made by the environment's composers 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 that 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 applications, we model the evental condition or *state* of each person (player) as a finite-dimensional topological vector space Γ_p of possible metaphorical states. This descriptive state space Γ_p is one factor of the complete state space M_p for each player, the other factor being the space S_p of possible sensed activity for that player within the environment, the space of sensor data which the system perceives. So

$$M_p = \Gamma_p \times S_p, \tag{7.1}$$

the space of all possible vectors of data the system maintains for a player p at any instant, is represented by a point in Γ_p (the metaphorical state), a point in S_p (for the sensor data), and their associated time derivatives. The complete space of states is the product of those describing each player:

$$M = \prod_{p=0}^{N_p} M_p = \prod_{p=0}^{N_p} \Gamma_p \times S_p \tag{7.2}$$

where N_p is the number of players. The original implementation of this system was built for five players, but the model has no limit. I prefer to work with three or more players to break up conventionalized dyadic interaction.

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

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. Figure 7.2 shows a sample diagram from the preliminary player state topology hand sketched by the composers of the TGarden 2001 system.

Metaphorical State Space

The base space Γ_p defines the possible metaphorical states that the player p in the interactive environment may inhabit. A configuration of potential states for the room is shown in figure 7.5; again, we think of the room in this context as another player. In our typical applications, the human player descriptive spaces are all identical, and one may think of the corresponding states as several points on a common space of potential human physical conditions.

When Γ_p has more than one connected component,

$$\Gamma_p = \Gamma_p^{(1)} \cup \Gamma_p^{(2)} \cup \cdots, \tag{7.3}$$

the player state $\gamma_p(t)$ is given by a point on each $\gamma_p = (\gamma_p^{(1)}, \gamma_p^{(2)}, \cdots)$. In this model, the substates on each connected component evolve independently of one another, and so 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 what follows.

Each space Γ_p of metaphorical states is built up as follows. The composers choose a collection of N fundamental states n_i, $i = 1 \ldots n$, each named after an elementary condition or scene such as Intrude, Feed, or Reveal. In the player state topology (figure 7.5), 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 that player is occupying,

$$\gamma_p(t) = \sum_{j=1}^{N} \lambda_j(t) v_j, \tag{7.4}$$

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

$$\sum_{j=1}^{N} \lambda_j(t) \equiv 1. \tag{7.5}$$

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

At any moment, a player inhabits a mixture of a particular set of “component” states, so Γ_p is restricted to certain permissible combinations, or *simplices* $\sigma^n_{jkl\ldots}$, where

$$\sigma^n_{jkl\ldots} = (v_j, v_k, v_l, \ldots) \tag{7.6}$$

is a simplex spanning those neighboring or component states. Think of a simplex as a span of vertices in some Euclidean space. If it spans three vertices then the simplex is a two-dimensional triangle; if it spans four vertices, then the simplex is a three-dimensional block (the tetrahedron whose body is the interior of the span of these four vertices), and so forth for any higher dimension. A player state may occupy a positive mixture:

$$\lambda_j(t) > 0, \textit{multi-index } j = k_1, k_2, \ldots k_n \tag{7.7}$$

only if the corresponding simplex $\sigma_{k_1 k_2 \ldots k_n}$ is contained in Γ_p. At any moment, 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.

The composers or designers of the responsive environment specify the set of fundamental states n_i and metaphorical associations to each, with cooccurrence relationships determining exactly which mixtures of fundamental states, or simplices $\sigma_{jkl\ldots}$, exist in the system.[18]

Transitions can be defined at simplex boundaries. 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 to allow a mechanism for mediating situations in which more than two simplices share the given boundary, and make it directed in case composers desire asymmetric transition relations between neighboring simplices. The weighted, directed graph g is represented by a matrix of floating-point values. Each entry $g\left(\{\sigma_i, \sigma_j\}\right)$ 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 consists of the polyhedron of a simplicial complex built out of vertex representatives of the states of the system, and determining which of those states may be simultaneously active, combined with the set of boundary relations $g(\{s\})$, 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 time $t \in [t_1, t_1]$ is a path in Γ_p determined by a dynamic that we describe in the sections that follow.

Sensor Space

As indicated earlier, the player state space M_p is made up of the descriptive space Γ_p outlined above, over each point of which lies a space S_p of possible sensor feedback that the system has about player p. The sensor data is used to drive the dynamics, in the manner described below, in response to the movements of the players within the environment. The sensor 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. Typical applications use very simple derivative features, 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 that represents the movements at a time 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 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

$$\Gamma_{\text{common}} \cong \Gamma_1 \cong \Gamma_2 \cong \cdots . \tag{7.8}$$

In cases where the human player states respond to similar sensor data, we have sometimes found it useful to think of there existing 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 of Γ_{common} with a uniform player sensor space S_p.

Energetic Model

For each player p_k, we engineer an evolving energy landscape over the state topology, letting the player state move as a massive particle on Γ_p, which evolves according to the laws of classical mechanics.[19] It is this movement that is recapitulated in the changing character of the output and responsiveness 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_p and sensor data vector $s \equiv s_p(t)$ of S_p is associated an energy given by:

$$U[\gamma] = \sum_{v_k \in \sigma(\gamma)} H_\Gamma(\gamma, v_i)\, H_S(s), \tag{7.9}$$

a sum over the sites v_i 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_i, while H_S is designed to give the energy of the player sensor data vector s relative to data assigned at v_i. A wide variety of energetic models based on equation 7.9 are possible. Those which we have implemented consist of a sum over pairs of factors having the general form

$$U[\gamma] = \sum_{v_k \in \sigma(\gamma)} f(\lambda_k)\left\{e^{-\beta E_k[s] + g_S \phi_k} + g_V V_k\right\}, \tag{7.10}$$

where $f(\lambda_k)$ is some function of the weights of γ_p relative to the states v_k. We describe below in more detail the Γ- and S-dependent contributions that we have found useful.

Figure 7.6 illustrates the simplicial model of state structure and parameters. The simplicial model is constructed from the data indicated. $\gamma_p[t]$ is the current state. $s_p[t]$ is the player sensor data vector. The data associated with each elementary state are the nominal sensor value μ, the variance, the static potential V, and the local scale of the sensor contribution to energy ϕ.

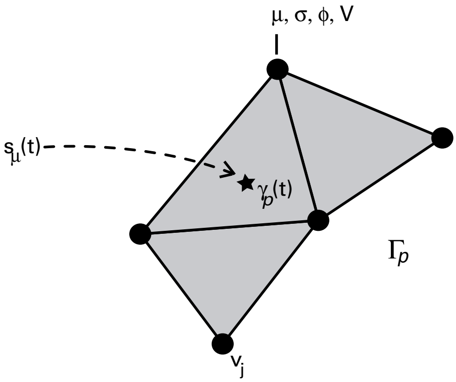

Figure 7.6
Simplicial model of state. Diagram courtesy of Yon Visell.

Sensor Coupling, H_S

Our sensor energetic coupling for the kth player has the form:

$$H_S(s) = \exp(-\beta E_k[s] + g_s \phi_k) + g_V V_k. \tag{7.11}$$

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.

$E_k[s]$ is the energy of the kth player's sensor data vector s relative to data at v_k. And β is a coupling constant for the model dependence on E_k, controlling how sensitive the system is to the player's sensor activity.

The local sensor coupling contribution E_k is determined in conjunction with a model parameter assignment procedure performed at the time the environmental dynamics are being designed. The composer assigns a nominal vector of sensor values μ_{ak} and variances σ_{ak} which are chosen to correspond with the metaphorical description of state v_k. (In case the player spaces Γ_p are identical, these data are typically chosen to be the same for all players.) Then we take a quadratic dependence

$$E[s] \sim (\mu_{\alpha k} - s_\alpha)^2 \tag{7.12}$$

because it represents the leading-order dependence near a generic potential with a minimum at the given sensor mean:

$$E_k[s] = \sum_{\alpha\beta=1}^{N_s} (\mu_{\alpha k} - s)(\mu_{k\beta} - s_{\beta,p}(t)) / \sigma_{\alpha\beta}, \tag{7.13}$$

where N_S is the dimension of the space of sensors S_p.

The mean sensor vector μ_α represents a set of sensor values which, to the extent that they agree with the player's activity, are meant to draw the player state $\gamma_p[t]$ toward the location where they are assigned. We have typically taken the variance $\sigma_{\alpha\beta}$ to be purely diagonal. Choosing values for the variance s allows the composer to assign the relative sensitivity of each state to different sensor vector components (sensor channels). A variance which is large relative to the real variation seen in a channel will leave the state relatively insensitive to sensor data in that channel.

If one desires to associate more than one representative nominal sensor vector and variance per site, one can have a weighted sum of such contributions to E_k. Because the generalization is not difficult, we restrict discussion here to the single summand case.

The Γ-dependent contribution to the energy at a player state 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_i on the simplex, and refer the standard embedding of the N-simplex in $\mathbb{R}^N$ as the constraint surface

$$\sum_{i=1}^{N} \lambda_k = 1\,, \quad 0 \le \lambda_k \le 1. \tag{7.14}$$

The constraint of convexity means that the N-dimensional components of any vector fields on the simplex should sum to zero, in order that they remain tangent to it.

To determine which potentials may be promising, it is useful to consider the force each would give rise to, using the formula

$$F = -\nabla U \tag{7.15}$$

(the dynamical results of which are explored in the next section). If we want the force to be continuous across simplex boundaries, we can decide that the normal component should vanish along each. One could also consider configurations in which this normal component attains a constant value on each boundary, and assign these boundary forces in a consistent manner, but the zero case is certainly the simplest, because it requires no intersimplex gluing considerations. The consistent application of gluing conditions can be obstructed by the topology of the model.

One solution is to take the force due to v_k to be polynomial in the coefficients λ_k and independent of the other coefficients:

$$F_k = \sum_{m>0} C_m \lambda_k^m \tag{7.16}$$

for some constants C_m.

The requirement that the force F be tangent to the simplex at each point, or in other words that its components F_k sum to zero, can be assured everywhere on the simplex in the simplest case by setting $F_k = C(\lambda_k - 1/N)$, where N is the number of states in the simplex. Such a force is obtained from an inverse potential,

$$H_\Gamma(\gamma, v_k) = -C(\lambda_k^2 - \lambda_k / N). \tag{7.17}$$

If the H_S term were location-independent, one would have a combined simplex potential given by

$$U[\gamma] \sim \lambda(t) \cdot \lambda(t) - \frac{\lambda(t) \cdot \mathbf{n}_s}{N} \tag{7.18}$$

with $\mathbf{n}_s = (1, 1, \ldots, 1)$ denoting the normal vector to the simplex, and $\lambda[t]$ the vector of weights for γ. The latter term can be thought of as a Lagrange multiplier enforcing the constraint of tangency.

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

$$H_\Gamma(\gamma, v_k) = C(1 - \lambda_k)^2, \tag{7.19}$$

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 $\lambda_k[t] = 1$, 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 $\mathbb{R}^N$.

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

$$d_{\mathrm{EU}}(\gamma, v_k) = \sqrt{\sum_j (e_k - \lambda_j)^2}, \tag{7.20}$$

where $\mathbf{e}_k$ is the unit vector, which is 1 in the kth component and 0 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

$$H_\Gamma(\gamma, v_k) = C\, d_{\mathrm{EU}}^p \tag{7.21}$$

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. The force, and consequently the momentum, have such discontinuities, but the player state trajectory is continuous. We will describe the dynamics further in the next section.

The environment can simply tend to drift to a steady-state local minimum, thus providing a closure to an experience corresponding to narrative closure in a conventional interactive scenario. And in the simplest case, time-based behavior can be incorporated through time dependence of the parameters of the system, which might be accomplished by treating a clock as a channel of virtual sensor data.

The pseudo-physical model, like any physical model, evolves by minimizing energy. In order to see explicitly how a state evolves, we treat it as a particle with an associated mass and compute the force acting on that state. We do this in the next section.

Dynamics

In this section, we describe how the model evolves the player states in response to the player sensor statistics. Evolution proceeds by integrating the first-order equations of motion (equation 7.15 and equation 7.16), 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,

$$F[\gamma] = -\nabla U - \xi\dot{\gamma}_p(t), \tag{7.22}$$

where we include a damping force depending on a coefficient x_i, 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

$$F_k = \frac{-dU}{d\lambda_k} - \xi\dot{\lambda}_k. \tag{7.23}$$

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

$$F_k(t) = (\lambda_k(t) - 1/N)e^{-\beta E_k[s_\alpha(t)]} - \xi\dot{\lambda}_k(t), \tag{7.24}$$

with E_k computed as in equation 7.13.

We then take F_k^T, the force tangent to the simplex, obtained by projecting out the (1, 1, 1, . . . , 1) component (perpendicular to the simplex). 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:

$$\dot{\lambda}_k \to \dot{\lambda}_k + \frac{1}{m}F_k^T(t)dt \tag{7.25}$$

$$\lambda_k \to \lambda_k + \dot{\lambda}_k dt, \tag{7.26}$$

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 composer 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 x, the coupling constants g_s, g_V of equation 7.11, and the sensor sensitivity β.

When channel-specific sensor sensitivity is required to be adjusted in real time, one factors from each diagonal component of the variance vector

$$\sigma_\alpha = \sigma_\alpha^{\text{mod}}\ \sigma_\alpha^{\text{static}} \tag{7.27}$$

a global factor $\sigma_\alpha^{\text{mod}}$ which may be varied, and likewise with the other local parameters of the system.

The palpable magnitude of certain of these parameters is constrained by the desired time scale for evolution for the autonomous and nonautonomous motion on Γ_p. The potential energy is the composer-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

$$T \sim \sqrt{\frac{m}{g_V V_k}}. \tag{7.28}$$

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 time scale for that response has the desired character, and it is easier to do so empirically once 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. This, however, is very sensitive to small variations.

One key point: These equations are used to implement a dynamical system in the computer, but the topological dynamics can be defined *without* explicit numerical constants and differential equations.

Experimental Work

The early exercises, studies, and installation-events by Sponge[20] dealt with particular questions in performance research: how to make events that were experientially as powerful as works of avant-garde theater but without resorting to verbal/written language, erasing the distinction between actor and spectator, and relying on thick, physical/computational ambient media. TG2001 as built by FoAM and Sponge was an installation-event that marked a transition and a bifurcation from performance research into a strand of public installation-events and a strand of studio-laboratory research in the Topological Media Lab (TML). After leaving Stanford for Georgia Tech in 2001, I started the TML to take stock of, and strategically extend, some of the technologies of performance according to a particular set of ethical-aesthetic heuristics inspired by continuity, human performance (e.g., the violin), human play (e.g., in water and sand), and nonelectronic matter like clay, smoke, or rain. I wanted to make responsive media synthesis engines and gestural instruments, and choreography systems that would allow participants to experimentally costructure, not "interact" with, coevolving ambient life in the "real time" of perceptually concurrent action and the specious present.[21]

Methodology versus Experimental Phenomenology

One of the first questions I was asked by colleagues from the sciences was: How do you know that this works? How do you measure progress in this domain? What is your methodology?

A negative response would be: "This is an activity incommensurate with your domain and your norms." It is difficult to fit methodologies for measuring utility-oriented applications to the composition and evaluation of expressive technologies, especially those based on statistical survey methods. As Barbara Hendricks said about the design of playgrounds: there is no average user.[22]

More constructively, we can observe that the domains of performing arts, installations art, and cinema all have their own systems of evaluation, mediated as much by the opinion of expert nonpractitioners as by popularity measures like attendance and market measures like box office receipts. The role of the critic, one who is informed by the history and conceptual ambitions peculiar to the context of practice surrounding the object of criticism, is well established with both practitioners and audiences. Even if that critic's judgment is contested, so long as the critic's responses are relatively coherent and stable, one can use them to triangulate the position and value of a piece of art with respect to one's own criteria. Peter Brook, a celebrated living director whose career spanned the Royal Shakespeare Theatre, Artaud, the Theater of Cruelty, and Grotowski, argued for the essential role of critics in the healthy ecology of theater.[23] Such a community of cultural beacons is well established as an evaluative network at least as accepted as a jury system composed of so-called peers. One problem with peer review is that such a system tends to average against innovations, perturbations far from equilibrium. One practical task in an emerging community of art practice could be to establish a critical practice alongside the material practice.

Human subjects committees have been designed to prevent abuses of subjects in experiments in psychology and medicine, influenced by controversies in the wake of World War II like Stanley Milgram's experiments in obedience to authority.[24] But then do such committees comprise the creeping edge of a Foucauldian discipline emergent among institutional experimentalists made docile? How else would one design phenomenological (not statistical or empirical) experiments but in the mode of art? However, to be interested in the phenomenological is not the same as to be a phenomenologist (i.e., an adherent of Merleau-Ponty, Heidegger, or Husserl), which would be a scholastic specialization. To be interested in the phenomenological is to be interested in the relational aspects of experience in embodiment and in the essences as they emerge in the course of lived experience. And most importantly it is to be sensitive to situation or context and how they may be bracketed. Human subjects research committees are an example of local rationality that masks global irrationality, just as "health" as a category in the market of commoditized medical care and risk management can ignore actual human well-being. Meant to protect experimental participants from being subjected to unethical procedures, "institutional review boards" or "human subjects committees" cannot govern a priori the essentially boundlessly unpredictable processes of creative invention in theater, choreography, or art making in general. When such committees do try to govern creative work, they must permit an unconditioned range of action adequate for creative design, or enforced rule-governed accountability tends to yield uninspiring work, or there is a disconnect between what is reported and what is actually done. In any case, this is clearly a situation of governmentality in institutionalized cultural production.

Technical Research

I approached the branching family of play spaces represented by TGarden, txOom, tgvu, and trg as phenomenological experiments of a certain kind, as events based on gesture and movement rather than on language, for people face to face in a common place, playing and improvising meaningful microrelations without language, in thick responsive media. I saw those play spaces as opportunities for ethico-aesthetic play, to adapt Guattari's concept of the coming into formation of subjectivity; as places to engage in biopolitics, radically dispersed into tissue and molecular strata, and reaching far beyond the computational media arts, meeting with experimental impulses in dance, movement, textiles, musical performance, experimental theater, but also the most speculative initiatives in urban design.

In his epochal essay on a revolutionary Theater of Cruelty, Antonin Artaud listed a mise-en-scène of theater in which all the theater arts—costumes, scenography, lighting, sound, body movement—would act in concert and with agency equal in power to the dramatic text. The Topological Media Lab has created its responsive media and instruments as conscious extensions of an Artaudian spectrum of technologies of performance.

However, the TML's broader aspiration is to create apparatuses for conducting speculative, critical, social, cultural, phenomenological experiments. Terry Winograd once commented that a phenomenological approach (versus human-computer interaction's nonrigorous appeal to the statistical empirical) draws insights from experiments of one, from singular experiments.

The ambition here is to conduct even the most philosophical speculation by articulating matter in poetic motion, whose aesthetic meaning and symbolic power are felt as much as perceived.[25] It is one thing to do philosophy or science *of* dance, and quite another to do philosophy informed by movement. And it is a further challenge to articulate movement-informed philosophy not in written text but *in movement* or some

Figure 7.7
IL Y A showing body gesture related to smoke, 2012.

other temporal medium. In order to prepare the ground for such a challenge, I shift the emphasis from representation to speculative, corporeal experience; hence the Topological Media Lab's emphasis on technologies of performance, live event, and real-time responsive media—on *alchemical matter*.

Alchemical Matter

To let people play immersed in media with evocative, responsive qualities, we could have them step into a warm pool of water laced with honey, so why use computational media? Computing the quasi-physics allows the creators to inject a physics that changes according to activity and local history, and responds in ways that resemble yet are unlike any ordinary matter. This is analogous to the alienation effect of theater, but not at the level of whole bodies (characters, actors, spectators, plot). Instead, what continuous, dense, topological dynamical systems afford is a microfine alienation effect at the level of substrate media such as calligraphic video, gestural sound, and kinetic fabrics imbued with uncanny physics.

Indeed it would take a lot of work to build up to macroscopic objects and actions from relatively homogeneous textures and simple dynamics. However, I would say that it is no more difficult or complicated than the enormous amount of hard-earned psychosocial knowledge, narrative apparatus, and literary skill needed to render a character in a novel or play from the raw material of alphabetic text and grammar. Such textural, alchemical techniques seem strange and unidiomatic for all of us who have been trained to the aesthetics and logic of whole bodies and macroscopic human-scale objects like words, props, characters, and theatrical or game action.

The Ozone state engine evolves through a rather sparse topological landscape with few valleys and peaks, whereas the visual and sound fields are synthesized as densely and temporally finely as possible, and as necessary to sustain a rich experience with millisecond-scale dynamic response that we do not attempt to compute using the slower state engine. The reason for decoupling the dynamical metaphorical state engine from the media instruments is in fact to decouple the evolution of the behavioral response "climate" from the dynamics of the visual and sonic textures, which should be as rich and tangibly responsive to the players' actions as possible. It seems artistically and compositionally useful to keep these dynamics decoupled from one another.

My concern in the context of this chapter is precisely with the possibilities that a microphenomenology—free of ego and anthropocentrism and indeed free of fixed, a priori objects—can offer toward fresh and refreshing improvised play. Aesthetically, this play should take place immanently in as dense an ambient medium as that of ordinary life. So the best approach would be to start with ordinary matter and real fleshy people in common space, and judiciously augment the everyday matter with just enough computational matter to give the event a strange and marvelous cast. This approach, which I nickname "minimax" design (maximum experiential impact for minimum computational technology), resonates with Grotowskian poor theater's choice of a minimalist technology of mise-en-scène relative to cinema, a minimalism which in fact is constitutive of its magic.[26]

The apparent inefficiency of such highly engineered virtual reality environments is in fact endemic not only to "bottom-up" simulations but to all simulations. As Maturana and Varela pointed out, to be as dense as life, a simulation of an autopoietic system can never operate any faster than that autopoietic system, and can at best run at the speed of life. So much for the cybernetic fantasy of mastering and replacing the lifeworld by a transcendental, superior simulation of life.

Rhythm as Structured Light = Calligraphic Video + Theatrical Lighting + Domestic Lighting

Classically, typography differs from calligraphy in two important respects. In a typographic alphabet, every instance of a letter form is mechanically identical, whereas a manually lettered form varies with every instance: identity is a limit rather than the instance. Also, typography is mechanically produced, whereas calligraphy is manually performed, and in fact breathed. Although it is not often taught explicitly, a calligrapher learns to time the rhythm of her breathing to the strokes, and so the pace and meter of her breath echo in the inked letter forms.

Calligraphic video refers to computationally synthesized moving image that fluidly responds to gesture. I suggested that video should be taken out of the screen, out of the box, treated not as a framed object but as structured light projected onto physical bodies. Different species of structured light have distinct textures and latent responses to perturbation. My approach was to borrow from computational physics and treat each frame of video as the initial data for a lattice computation approximating the dynamical equations for some physical substances. For example, Yoichiro Serita carried out this idea treating video with the heat and wave operators. (See figure 7.7, showing calligraphic video as used in the IL Y A video membrane.)

Using these computational lattices that numerically simulate the dynamical equations of physical materials imbues the video with analogous responsive qualities. Projecting such video textures allows people bathing in such structured light to nuance the video textures as if they were physical material. This leverages the deeply embedded intuition that we lay down from infancy with the materials of the world, an intuition that may not be articulated in any language. Moreover, this lattice computation scales well with increasing numbers of people. In fact the computational complexity is O(number of lattice cells), but is essentially a constant with respect to the number of people in the field represented by the lattice. This is the same strategy followed by astronomers, who model galaxies with tens and hundreds of thousands of stars by smoothing the distribution of stars into a continuous distribution density, and then model the physical motion by hydrodynamical equations of a plenum rather than the combinatorially intractable, multidimensional graph of pairwise-interacting particles.

A low-resolution display is a set of isolated "pixels" spatially scattered across a space in some arbitrary, though sometimes intentionally determined, pattern. These so-called pixels could be the set of all the LEDs winking in your house, or a set of ceiling fans, or sewn in strands of glowing fiber sewn into a carpet.

What are needed are abstractions to operate these low-resolution displays, not as regular rectangular arrays of a million pixels, but amorphous, scattered, and shifting sets of isolated flecks of light, puffs in air, or curling polymer. Whatever the medium, these sets can be articulated more expressively by the extensive concepts and operations of measure theory than by classical geometrical concepts and operators. The richest descriptions we have of such point sets draw not from geometry (whether Riemannian or otherwise) but from measure theory. There exists computational support for this at a primitive level, but not yet incorporating the more expressive mathematical articulations that we would need.[27]

Rhythm as Gestural Sound

Similarly, sound can be *nuanced*, rather than triggered as playbacks of fixed recordings by the movement of a body or a thing. Typically, in an "interactive" installation, a specific movement by the human triggers the playback of a recorded sound. Even though the sensing of the movement may be a sweep of the hand through a laser beam or some more exotic mediation, the result is still limited to the playback of a piece of sound, whose audible quality does not vary in the course of the playback from the way it was edited. This treats sound as an object—a compact body in naive space-time, which is often how a visually oriented artist first approaches the use of sound. Yet another symptom of a visualized and object-oriented approach to sound is heavyweight sound "spatialization" systems that focus on *placing* sound objects in a virtual three-dimensional space, or flying them along spatial trajectories as if they were physical compact bodies rather than rhythms permeating a continuous material field. However, a sound that is synthesized or varied over the course of its running time according to some parameters fed by the nuance of a live, coordinating movement can be gesturally modulated. Think of a piano versus a violin. After the piano's hammer falls, the note resounds in a way that is pretty much independent of how the pianist lifts her or his hand after removing it from the key. Because the violinist maintains continuous contact with the wood and strings and bow, however, she or he continuously nuances the note as it is sounded. This continuity of contact with gesture and consequent continuity of the nuancing of the sound characterizes what I call gestural sound, as contrasted with triggered sound. Gestural sound is analogous to calligraphic video.[28]

One approach to creating electroacoustic instruments is simply to electronically mimic the analog instrument, with momentarily amusing results. Another is to mime the forms of performance, for example to create an "opera" or a "symphony" or a rap performance, but with electronic instruments in place of conventional ones. While this may comfort those who need to measure progress by triangulating from shore-bound landmarks, these imitations strongly bind the creation of fresh sonic experience. Fifty years after John Cage oriented us to the experience of "organized sound," we need not simply repeat the instruments or the genres of sonic practice as we explore acoustically mediated experience, except as exercises to build confidence in technique.

My approach has been to treat sound as a medium for exploring distributed agency, superposition, and alinguistic time-saturated patterns. Creating "instruments" that map movement to sound permits my collaborators and me to experiment with

Figure 7.8
Navid Navab at a sonified table, Topological Media Lab. Photo by Jérôme Delapierre, 2012.

intentionality, agency, and affect via gesture. Concretely this has demanded creating sound synthesis instruments that are not mimetic of any conventional instrument. Why only imitate something that is already done quite well and far more subtly, sustained by centuries of performance practice and social embedding? Thinking of sound as substrate instead of well-schematized patterns implies making *sounds* rather than *music* or *speech*. Even sonic patterns that happen to be those of speech or music can be treated as texture rather than semantic representations, with fresh results.[29]

Rhythm as Movement: Dance, Movement Art

As Arakawa and Gins proposed in their provocative book *Architectural Body*, the body finds its limits where attention alights. We can pose this more generally as a phenomenological problem of how the body is conjured in the wearing of a space's material.

Summary

Across all these streams of inquiry, we study how to richly condition streams of media in concert with the activities of the people in a common built environment, such that the activities (to use an overly anthropocentric term) of the media and the people costructure one another and evolve over time according to prearranged strategies and latent predilections, contingent activity, and memory of past activity. I appeal to continuous dynamical systems on several grounds:

(1) People's experience of the world is continuous.
(2) People have sedimented huge amounts of experience with the physical world, so we should leverage it by using quasi-physics models.
(3) I wished to see how we could move away from egocentric and anthropocentric design.

The Atelier-Lab as a Transversal Machine

On the macro scale of sociodynamics and cultural-epistemic diffusion, my long-term interest in the TGarden and its sibling responsive play spaces extends beyond the actual events themselves to the mixing of ideas and conflicting ideological commitments from different epistemic cultures. It could be liberating to practice our arts and sciences in a more reflexive way.[30]

Two decades ago, Félix Guattari pointed to the heterogeneous machines around us: material, semiotic/diagrammatic/algorithmic, corporeal, mental/representational/informatic, libidinal/affective. Guattari's *Chaosmosis* asked how we could construct machines that act transversally across those machines. In the decade since 2001, the Topological Media Lab[31] has been working as an atelier-laboratory transversal to computer science, performing arts, and architecture and the built environment, to generate insights and techniques in the domain of new media and responsive environments. The atelier was motivated by the questions: How can ordinary actions in everyday environments acquire symbolic charge? What makes some environments enlivening and others deadening? Reflexively, we ask: To what extent can we instantiate labs or ateliers for the creation of apparatuses for ethico-aesthetic improvisation? This section describes institutional, sociotechnical, political and economic issues around running such an atelier-laboratory as an alternative social economy complementary to postindustrial, "knowledge-based" economies.

The big methodological moves are (1) avoiding a priori schema, (2) working with material, collective, environmental situations, and (3) moving from nouns to verbs, from things to transformations of things. This includes, by reflexivity, moving from working with fixed (i.e., transcendentalized) concepts to putting concepts in play. It motivates an approach via ontogenesis. Staying close to the material and collective-environmental implies making "thick" experiments in the "wild."

The TML prototypes what I call an atelier-lab, an open space in which affiliates can pursue art research without having to constantly defend individual projects in the institutional language of disciplines and granting agencies, or in terms of the market.