Calligraphic Video: A Phenomenological Approach to Dense Visual Interaction

Sha Xin Wei
Topological Media Lab
Concordia University
Montreal, QC, Canada
xinwei@mindspring.com

Michael Fortin
Topological Media Lab
Concordia University
Montreal, QC, Canada
michael.fortin@gmail.com

Jean-Sébastien Rousseau
Topological Media Lab
Concordia University
Montreal, QC, Canada
jsrousseau@gmail.com

ABSTRACT

No matter how the image is computationally produced, screen-based graphics are still typically presented on a two-dimensional surface like a screen, wall, or electronic paper. The limit of manipulating objects in a two-dimensional graphical display is where each pixel is an independent object. These two observations motivate the development of *calligraphic video*, textures that can be manipulated by the user using intuitions about physical material such as water, ink, or smoke. We argue for a phenomenological approach to complex visual interaction based on corporeal kinesthetic intuition, and provide an effective way to provide such texture-based interaction using computational physics. A motivating application is to create palpable, highly textured video that can be projected as structured light fields in responsive environments.

Categories and Subject Descriptors

J.5 [**Arts & Humanities**]: Arts, fine and performing, 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; J.7 [**Computers in Other Systems**]: Real time

General Terms

Experimentation, Human Factors, Algorithms

Keywords

responsive environment; tangible media; responsive media; computational physics; gesture-based interaction; realtime video; real-time sound; phenomenology; design

1. COMPLEX DISPLAYS IN THE LIMIT

Decision-making activity with complex displays impose a cognitive load that generally increases with complexity. The challenge of manipulating displays in their full complexity can overwhelm the human capacity for interpreting, remembering, and manipulating icons or object-based graphical structures. On the other hand, the experientially richest visual displays generally come from sampled video textures. Our basic conceptual contribution is to realize that richness comes from corporeal experience, and is independent of combinatorial structure, which is a formal property not amenable to intuitive grasp. We apply this to the problem of fully exploiting ubiquitous two-dimensional computer graphic interfaces leveraging intuitions derived from corporeal experience for use in applications to new performance as well as built-environments. (Figure 1.)

On the computational side, considerable computational resources are needed to generate 3D or higher dimensional models (in the case of particle dynamics or animated object physics) but ultimately the results must be projected back to a two-dimensional array for display. Object-oriented graphics algorithms must contend with combinatorial complexity as a function of the number of objects and of dimension. Our array-based computational physics techniques use (variably sized) lattices with various underlying data representations, but are robustly invariant in the number of objects and points of user manipulation because each pixel effectively can be a manipulable particle.

2. PRIOR WORK ON KINESTHETIC INTUITION

In 2001, Sha Xin Wei built with colleagues[1] a responsive environment in which wireless sensors beamed accelerometer data to OpenGL textures that were mapped onto a polygonal mesh. This textured mesh was projected onto the floor from a height of 20', which produced moving "wings" registered to body of the participant. (Figure 2.)

The mesh width varied according not to some pre-scripted clock-based logic but to a function of the contingent movement of the participants' bodies. Despite the crude graphics, jumping on the hard floor onto which this responsive mesh was projected, one felt as if one were jumping on an elastic rubber sheet. Sha concluded that what gave such a strong sense of elasticity to the projected mesh was the nearly zero latency synchrony of the mesh's grid size with the vertical forces on the participant's body, as measured by torso-mounted accelerometers. This motivated the strategy of using semantically shallow models that do not infer the cognitive, intentional or emotional state of the participant, but instead parameterize the physics driving the graphics animation, with perceptibly negligible latency. A major limitation of that early work was

[1] Sponge and FoAM art research groups, based in San Francisco and Brussels, respectively. http://www.topologicalmedialab.net/xinwei/sponge.org, http://f0.am

MM'09, October 19–24, 2009, Beijing, China.

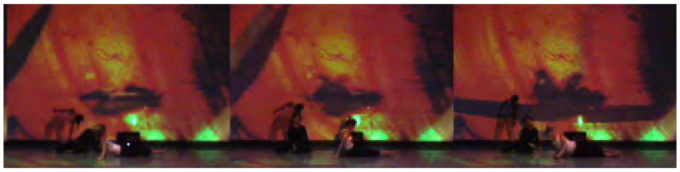

Figure 1: Ouija experiments on collective or intentional movement. Dancers under projection of live brushwork by F. Radonjik, composited with *calligraphic video* by JS. Rousseau synthesized from their own movement. (Montanaro, Sha et al., Topological Media Lab)

the coarse resolution of the 3D geometry that we could render and drive in real-time from sensor data. In subsequent work, we decided to strategically forgo 3D graphics and use the computational overhead to present much richer 2D textures.

Normal mapping, for example, uses pre-computed perturbed normals from a bump map stored within a texture. The advantage of this method is that it adds extra detail that reacts to light without altering geometry, very useful for adding small details such as reliefs on a textured surface. For example, [4] shows how a complex mesh can have the number of polygons reduced while maintaining detail.

If we were to take this simple technique divorcing it from its 3D context and apply it to a 2D surface covering the screen then the geometry would no longer be an issue. The rendered detail would be constrained just by the size of the source texture which can provide the richest amounts of detail possible as it reaches the limit case of the number of on-screen pixels.

We call physics-based processing of video textures that exhibit quasi-physical response to gestural manipulation, *calligraphic video*. This is a special case of *responsive media*, media that evolve, up to the limit of human perception, concurrently and determinstically with human gesture and movement. We are interested in *calligraphic video* for both engineering and scientific reasons. In practice, our screen-based visual interfaces can become so dense that hierarchical and visual-iconic representations saturate our human faculties. Our strategy is to tap the human'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, over a lifetime, so it seems reasonable to leverage that sort of lifelong, pre-verbal capacity. We emphasize that these applications complement and do not replace conventional visual interfaces based on object models or, formally, graphs of objects. Texture-based interfaces may also have value in manipulating very large numbers of time-based streams as well.

The other principal motivation for *calligraphic video* is to experimentally study the human construction of corporeal kinesthetic intuition via continuous gesture in continuous media. This is aligned with a phenomenological approach to embodied cognition that draws on the work of Edmund Husserl [12], Eugene Gendlin [9], and Humberto Maturana and Francisco Varela [17]. Gendlin's work reintroduced into psychology the phenomenological study of "how concepts (logical forms, ..., rules, algorithms,..., categories, patterns .....) relate to experiencing (situations, events, therapy, metaphoric language, practice, human intricacy ....)." (Gendlin [9], i) Biologists Maturana and Varela have introduced the study of organisms as *autopoietic systems.* Partly motivated by such considerations, in [7] Dourish argues for a phenomenologically informed approach to designing computational systems.

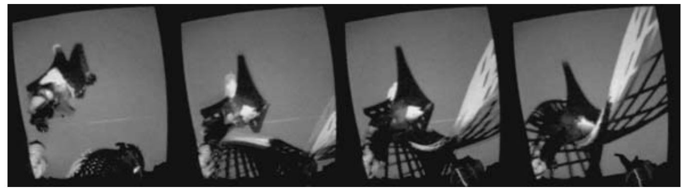

Figure 2: Participants in a responsive environment, under projected meshes that morph according to their movement. TGarden, Ars Electronica, Linz 2001

2.1 Kinesthetic Intuition Complements Cognitive Work

In the limit case where the number of manipulable objects approaches the number of pixels in the display, we introduce a kinesthetic approach to video-based interfaces that also leverages the physical intuitions of continuously manipulating continuous matter. Using a phenomenological approach may sidestep the cognitive load imposed by complex visual displays.

Entertainment applications such as gesture-based games with video feedback rely on sensuous-aesthetic as well as cognitive decision-making activity. Our approach is one approach to simultaneously aesthetic and sensuous intuition arising from prior reading of designed graphics, and from the body's tacit, expert, lifelong negotiation of the physical environment. The sort of intuition we leverage is that of *continuous* matter and *continuous* dynamics.

2.2 Phenomenology and Gesture

We are building these *calligraphic video* interfaces as platforms for research in gesture along phenomenological lines. (Sha [21]) Gendlin complements the logical structure of cognition with *felt meaning*, which has a precise structure: "Experiencing is 'non-numerical' 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." (Gendlin [9], v). Moreover, categories may be logically but not experientially prior to instances.[2] 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 pre-existing category. For example, graphic objects in our manipulable system must not appear to the user as built out of a pre-existing 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 – *morphogenesis*, to use a term of René Thom (Thom [24], [25]; Petitot [19]). We build *calligraphic video* – video texture that responds to manipulation by human gesture as interpreted from camera-based input – as appa-

[2]Gendlin stands in for a vast literature of related work in psychology and phenomenology, including Leibniz, Husserl, Merleau-Ponty, Peirce, and James. But for compactness we refer to a few bridging scientists.

ratuses[3] 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 pre-existing logical, functional, or geometric categories). The video texture substrate may not appear uniform at all, but it is continuous in space and time. Now, rather than use arbitrary dynamical systems to animate the responsive video, we choose to study the structure of corporeal-kinesthetic-visual intuition[4] via improvised manipulations of media that leverage corporeal-kinesthetic-visual experience of continuous matter commonly encountered from childhood.

3. COMPUTATIONAL PHYSICS

We describe three computational models of physical material that we have implemented for real-time video processing. These lattice models are based on the Laplace heat equation, the Navier-Stokes equation for turbulent fluids, and Ginzburg-Landau equation for magnetic domains. Most of these were implemented on conventional machine architectures. These three models cover a range of physicality that provided some sense of the phenomenal richness required and the limits of perception.

Over the past four years, we have successfully built a suite of real-time, array operators on video streams for each of these PDE's. Each PDE operator treats a frame of video as initial data and generates a new stream of arrays as short time evolution solutions of the corresponding PDE intercalated into the incoming video. The numerical simulation is intercalated in between frames of incoming video, balancing computational complexity, computation grid size, video resolution, and video i/o bandwidth. Put another way, each incoming frame of video is used to set instantaneous 'initial' conditions that re-trigger the evolution of the PDE.

In the following three sections, we present the models proceeding from the simplest to the most sophisticated material model. In each section we present the model in compact physics language. Then we describe the computational implementation of these models for real-time video, in enough detail so that experimental apparatus may be evaluated.

3.1 Introduction to Lattice Methods: Heat

Simulating the diffusion of heat through a homogeneous medium provides the canonical and simplest physical model for a lattice computation. This initial data is integrated by our real-time implementations of these simulations in between frames of video, so these effects are realized and experienced concurrently with the activity of the performer or participant. If $\phi(\mathbf{x},t) : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ is a scalar field on spacetime

$$\frac{\partial \phi}{\partial t} = \triangle \phi \qquad (1)$$

This partial differential equation, known as the Laplace equation, can be approximated on a discrete rectangular lattice using finite differencing, and numerically integrated using a relaxation method. Where $C_{i,j}$ is the value of the (i,j)-th cell, the method in essence is given by (Figure 3):

[3]We interpret apparatus in Karen Barad's richer sense of a hybrid of matter, expectation and theory. See chapter 4 of Barad [2].

[4]Such experiments are an empirical approach to Husserlian studies of intuition and experience. See Husserl [12].

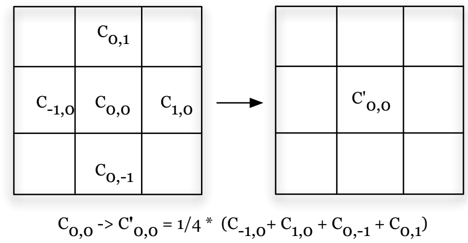

Figure 3: The center cell's value is replaced by the average of its neighboring cells' values.

$$C_{i,j} \to \frac{1}{4}(C_{i-1,j} + C_{i+1,j} + C_{i,j-1} + C_{i,j+1}) \qquad (2)$$

3.2 Matter Intuition: Navier-Stokes Model of Fluids

More recent results are based on the Navier-Stokes equation (See Chorin and Marsden [3]). A fluid whose pressure and temperature are nearly constant can be characterized by a velocity field $\mathbf{v} : \mathbb{R}^n \to \mathbb{R}^n$ and scalar pressure field $p : \mathbb{R}^n \to \mathbb{R}$, where $n = 2$ or 3. We can characterize the assumption that the fluid conserves mass by the condition that there are no sources or sinks in the velocity field. By Stokes theorem, this is equivalent to the condition that the velocity field is divergence-free:

$$\nabla \cdot \mathbf{v} = 0 \qquad (3)$$

Our second physical assumption is that the momentum is conserved. Let ρ be the density, p the scalar pressure, ν be the viscosity coefficient, $\mathbf{f}$ the external force. Newton's Second Law states that force is the rate of change over time of momentum:

$$F = \frac{d(m\mathbf{v})}{dt} \qquad (4)$$

Applying the chain rule for the total change of a function $\mathbf{u}(\mathbf{v}(\mathbf{x},t))$ of velocity field $\mathbf{v}$ which is in turn a function of position and time, one can show that the total derivative of $\mathbf{u}$ is given by:

$$\mathbf{D}[\mathbf{u}] = \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{u} \qquad (5)$$

Equation 5 defines the *material derivative* of $\mathbf{u}$, which yields the physical change of some material field $\mathbf{u}$ that is carried along by the velocity field $\mathbf{v}$ with respect to time. We can write the total force acting on an infinitesimal fluid element as:

$$\frac{\partial(\rho\mathbf{v})}{\partial t} + (\mathbf{v} \cdot \nabla)(\rho\mathbf{v}) = \nabla p + \nu \triangle(\rho\mathbf{v}) + \rho\mathbf{f} \qquad (6)$$

Solving for the time derivative of the velocity field, we obtain the Navier-Stokes equation (Chorin and Marsden [3]):

$$\frac{\partial \mathbf{v}}{\partial t} = -(\mathbf{v} \cdot \nabla)\mathbf{v} - \frac{1}{\rho}\nabla p + \nu \triangle \mathbf{v} + \mathbf{f} \qquad (7)$$

(There are many versions of this equation in the literature.) For the densities, we just deal with the first term, known as the *advection equation*. This moves the densities according to the velocity field. (Stam [23] includes a diffusion term for the fluid itself. We do not implement it.)

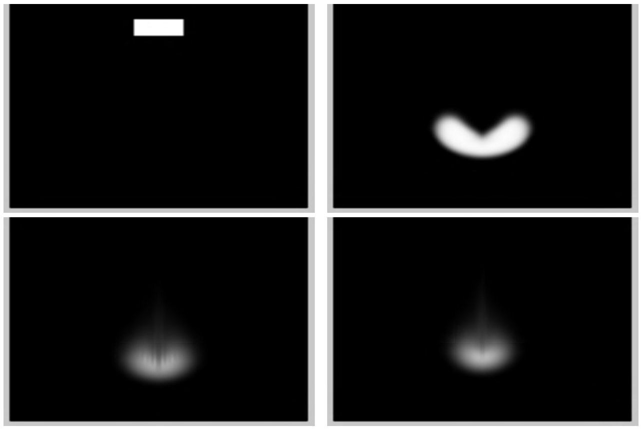

Figure 4: A test case used to test the implementation of the Navier-Stokes model. Top-left is the initial state of the test-case, gray walls with a white block that continually has a downward force applied to it. The force follows the block as it gets advected.

$$0 = \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{u} \tag{8}$$

3.3 Implementation of Navier-Stokes Model

To create freely parametrizable, real-time responsive video "fluids", we implemented the Navier-Stokes Equation (7) on multiple processing environments, notably the CPU and the GPU. The CPU version has been transformed into a plugin for the Max/MSP/Jitter software development platform. This section details our implementation efforts on both the CPU and GPU.

Our implementation is based on the GPU version of Stam's [23] unconditionally stable solver as implemented by Harris [11] extended using a MacCormack method for the advection as implemented in Crane *et al.* [5] for more accurate results.

Inspired by the *jit.repos* object from Cycling '74's Jitter framework [6], we advect a grid where each cell contains its current coordinates using the MacCormack method. Then for anything that we need to advect we would look in this grid, where each cell contains the coordinates of where to fetch data. Using this simple caching technique we are able to reduce the number of computations when advecting multiple fields and obtain a noticeable speed increase in the Jitter plugins.

The caching technique used is similar to the advection of texture coordinates Stam used in [23] to get richer detail out of the simulation. The difference is that we re-initialize our grid in each frame since it is used as an integral part of the advection phase to move densities and velocities. Given that texture coordinates are usually resolution independent, we allow the velocity field to be smaller than a given advected field to project a more high resolution version of the fluid while reducing the number of calculations.

The CPU version of our code uses pthreads as an underlying threading API and atomic instructions to synchronize the threads. The threads use an atomic instruction to increment a variable that corresponds to the currently processed row. To reduce the number of cache misses, we ensure that the data of each row to be processed has its elements consecutively stored in memory and access the data in sequence so that the hardware prefetcher will kick in and start loading data into memory.[5] We have not optimized advection. However we refer the reader to Kim [15] on optimizing Stam's unconditionally stable solver – a paper that was found after the implementation.

We broke the solver into independent Jitter plugins, allowing us to easily swap out individual components of the fluid solver and use them in varying contexts. Since we can work with the native data format of Jitter, it means that an application programmer can easily manipulate the underlying fields as they evolve through the simulation using Jitter's built-in functions.

The GPU version of our code, not yet ported to Jitter, uses 16-bit floating point color components.[6] On contemporary conventional machines, using 8-bits per component is problematic due to accuracy problems which results in the velocity field losing all momentum. For simplicity, we currently do boundary checks within the fragment program using conditional statements, however as suggested by Woolley [30], for more speed we can render different primitives for obstacles and free space.

Figure 5: Silhouettes of two people under Ising model texture

3.4 Limits of Physical Intuition: Magnetic Spin Models

We have extended this work to magnetic domains and consider Ising spin models in order to traverse and study the limits of embodied intuition. The Ising model (Figure 5) for magnetic domains – and its generalization from discrete range $\{0,1\}$ to continuous range of the circle $\mathbf{S}^1$, the Ginzburg-Landau model – provide quite different physical models for the continuous, real-time manipulation of a continuous texture:

$$\imath u_t^\varepsilon + \Delta u^\varepsilon = \frac{1}{\varepsilon^2}(|u^\varepsilon|^2 - 1)u^\varepsilon \tag{9}$$

in which vortices form as $\varepsilon \to 0$.

The simplest way to model this is to use the relaxation method as for the Laplace heat equation, but where the values are taken in $\mathbf{S}^1$ rather than $\mathbb{R}^1$. Further experiment is needed to determine whether these simulations of magnetic spin domains driving re-synthesis of the input video, can constitute substrates for intuitive manipulation of time-varying textures.

In any case, we now have at hand a variety of pseudo-physical textures with which to construct experimental "matter" for phenomenological experiments.

4. APPLICATIONS

Our applications primarily lie in the domain of live performance and associated phenomenological experiments. In June-July 2007 we conducted 4 weeks of a series of dance experiments called *Ouija*

[5]For CPU optimizing we refer the reader to Gerber et al. [10].

[6]RGBA - each component is 16bit, thus a 64bit vector to hold the color with alpha.

Figure 6: Visitor jumping rope in responsive environment transforms herself into 'fire' based on her own movement.

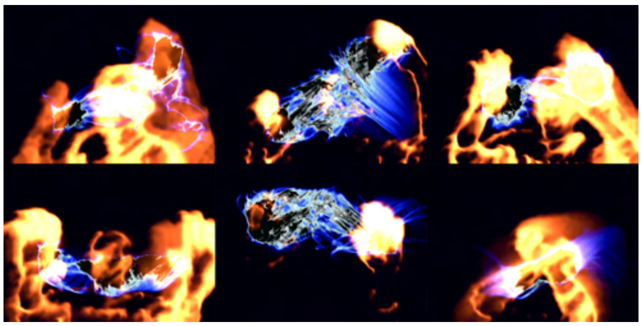

Figure 7: Six frames from improvised stretching of taffy material between two hands. We do not track blobs but treat optical flow as a lattice of gravitating matter.

in a theatrical blackbox at Concordia University. (See Figure 6 and Figure 1) This series of phenomenological experiments was designed with a choreographer Michael Montanaro[7] to see (1) how we could distinguish between intentional and non-intentional gesture, and (2) how such movement is modulated by the presence of responsive or pre-edited video and sound. The next subsections introduce some of the tools and methods we used during those experiments and look at other similar developments in real-time video involving performance.

We have implemented our video instruments, as well as the Ozone [22] media choreography framework in Max/MSP/Jitter (Max) a graphical programming environment for realtime media processing.[8] The programs and applications are coded using different frameworks: Max abstractions (sub-programs), Java or C externals, and OpenGL shaders (in jxs).

4.1 Fields, densities and lattices

Testing and developing applications for live performances and phenomenological experiments involving diverse, varying groups of participants notably pushed the development of scalable real-time visual instruments as part of our Ozone media choreography system [22].

Our tracking method of choice, following our general strategy of using field (lattice-based) methods wherever possible, has been the Lucas-Kanade optical flow method available from a standard computer vision library inside Jitter.[9] Instead of tracking discrete points, or barycenters of blobs, we track motion across the entire visual field, which gives continuous spatial information that is much more robust and general than blob tracking and provides dense data that we can stream into the rest of the video processing network. (See Figure 7.)

[7]Michael Montanaro, Chair of Contemporary Dance at Concordia University, designed the structured improvisation structure of the experiment. Assistant choreographer Soo-yeon Cho supervised the actual movement exercises. See video documents on the *Ouija Experiment On Collective Gesture In Responsive Media Spaces*, http://www.topologicalmedialab.net/joomla/main/content/view/159/74/lang,en/

[8]Max/MSP/Jitter version 5, 2008, http://www.cycling74.com [6].

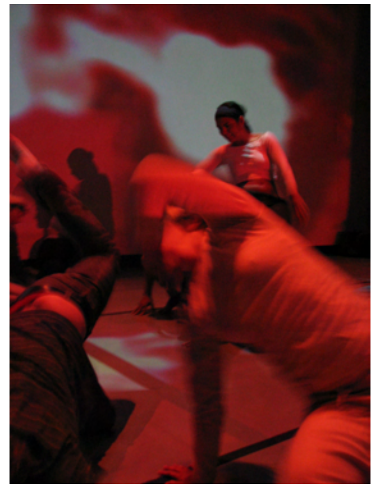

Figure 8: Ouija movement experiments. Dancers' movement parameterizes gravity and Navier-Stokes fluid models that in turn are projected as lightfields in place of conventional theatrical illumination.

4.2 Optical flow to fluids to painterly lightfields

A notable example of a responsive installation using continuous input to generate a graphical output is MeteorShower. (Figures 6 and 9) This instrument uses a 2D lattice of attractors to influence a set of freely moving particles. Being driven by a custom gravitational model, the particles circulate over the entire field of attractors. The attractor masses are scaled according to the motion gradient from the realtime camera feed. Scaling the masses of the attractors modifies the accelerations of the particles. Each particle evaluates the force exerted on it for every attractor of the lattice for each frame. Unlike a flocking algorithm, the moving particles do not interact, and therefore computation of the forces occurs only between particles and attractors. In Figure 7, for example, the player's hands' optical flow constitutes a field of gravitating matter, so he can stretch the cloud of particles like taffy with slew and momentum and elasticity. By varying quasi-physical constants like friction, elasticity, time-decay of forces, we can design a continuous range of "matter" with extremely diverse behavior. The Ouija experiments (Figures 1 and 8) investigated the impact of quasi-physical lightfields on the emergence of movement that was more or less intentional, and more or less collective. Over four weeks we staged systematically varied scenarios: with no media, with responsive sound, with responsive video-illumination, with both responsive sound and video fields; with prepared movers (dancers), with un-prepared movers together with prepared movers. We mapped movement into our quasi-physical models, and in turn projected the resulting video from overhead and side projectors into a 17m x 17m x 8m high black-box space in place of and together with conventional (DMX controlled) theatrical lighting. [10]

[9]cv.jit by Jean-Marc Pelletier, IAMAS.

[10]We report the results in a companion article (Sha and Montanaro,

Figure 9: MeteorShower installation. Visitor's optical flow provides mass density, and also the perturbative velocity fields.

Figure 10: Blending realtime video into paper to create composite temporal material for a chrysalis containing the dancer, in Touch2 performance, Remedios Terrarium, Montreal 2008.

4.3 Linking instruments together

One of our main goals has been to make it easy for composers to link video and sound instruments together and to orchestrate sets of instruments. The MeteorShower instruments, for example generate not only graphics, but also particle data, which we pack as a Jitter matrix streamed over the local network to sound instruments on a different instance of Max. A particle matrix can contain for example 12 planes of data: id, positions, velocities, accelerations, charge, mass. The heart of the challenge then becomes designing an effective, and interesting mapping. In the past we have successfully mapped particles positions and velocities to a number of "voices" using a Max/MSP granular synthesis instrument created by Tim Sutton, the principal responsive sound instrument designer at the Topological Media Lab.

4.4 Balancing CPU and GPU loads

Using Jitter's matrix array data structure the composer can quickly and efficiently manipulate an array of floating point or integer (char)

in preparation).

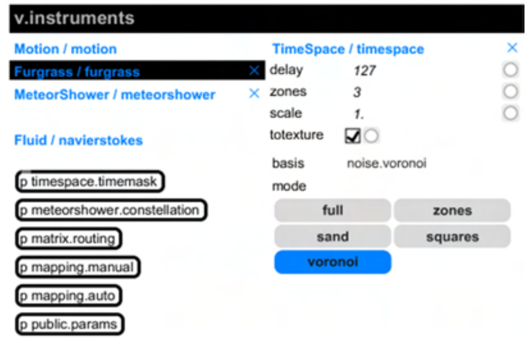

Figure 11: Chainable realtime video re-synthesis instruments and OpenGL processing in Max framework: Timespace pixel-level time-shifting, Meteorshower general particle system, Furgrass vectorfield, as well as Navier Stokes fluid simulation engines.

data, and visualize or process its planes as video images. Furthermore Max/Jitter provides facilities for acquiring and broadcasting such arrays as streams on networks between CPU's.

Another way to operate on 2D rectangular data like images and videos is to use OpenGL textures, slabs and shaders (jit.gl.texture, jit.gl.slab, jit.gl.shader), which are all available in Jitter. Once transferred to OpenGL textures, we can transfer computation to the GPU. Our instruments and other modules for real-time video provide a large range of abstractions that leverage the capabilities of the GPU without the need to code customized shaders, though we have written custom shaders as necessary. Numerous shaders are already part of the Max release, and others have been written for specific purposes. Combining live video, synthesis and GL shading modules, we readily synthesize video as structured light-fields to augment the movement of paper (Figure 10) or elastic membrane (Figure 12) in live performance. Figure 11 shows a typical front-end for a set of video instruments as an interface to the set of fluid externals described previously in this paper.[11]

Currently, all calculations executed prior rendering are done by the CPU, while the GPU is being used for rendering, shading and compositing. Most CPU calculations that are happening inside the visual instruments are done using Jitter matrices. While this is still not as fast as doing GPU hardware processing, doing operations on Jitter matrices is significantly faster than using expressions or *for / while* loops which incidentally do not exist in Max as native constructs. By leveraging Max's unique model, we are able to maintain real-time processing capabilities while providing simple interfaces for media designers and performers who wish to use our system.

5. CONCLUSIONS

5.1 Prior Work

In Section 2, we have surveyed theoretical prior work, and the arguments for the approach described in this paper. This section situates our approach with respect to prior work on user interfaces, graphics/ animation-based interfaces, and enactive interfaces.

For example, there a significant body of work on the use of physical models for interfaces based on graphic animations, distinguishing between "phenomenological" and "causal" (e.g. "spatio-temporal shape or force field") effects, where phenomenological

[11] See Touch 2 video of dancer http://www.topologicalmedialab.net/video/movement/touch_2_fx.mov.

is equated with 'mere' appearance to use for example Neyret & Praizelin [18]. It is important to observe that most of these computer graphic applications are primarily concerned with ocular inspection of some visual object, rather than kinesthetic, embodied experience ([28] for example). We use phenomenology in its original sense (Husserl [12], Gendlin [9]) to emphasize embodied, felt experience, distinct from visual representation. Loke et al. argued on similar phenomenological grounds that "movement is constituent of perception" ([16] 693). Their study of the Sony Eye Toy is restricted to reduced game schemas of *a priori* objects and movements. Rather than restrict our interaction to abstracted conversational turn-taking models, we treat events that are dense with *concurrent, continuous* movement by humans and media.

With respect to gesture tracking, our emphasis is not so much on sense-modality, as is commonly done in the experimental studies of enactive interfaces (see Visell [27] for an explication centered on visual-tactile modalities). Rather, we focus on what Husserl recognized as the *apperception* of an object. Furthermore, instead of representing and recognizing classes of gestures defined *a priori*, on which the pattern system is trained (Varona et al. [26], for example), we focus on creating visual and sonic fields with usable and intuitable quasi-physics.

Will.0.W1sp [29] is "an interactive installation using real-time particle systems to generate characters which move with human motion, but have no set form." However, Will.0.W1sp uses a mass-spring system model whose physics is much reduced. In that case, the motion capture data is manually processed via rotoscoping into virtual skeleton joints. Attractors are associated with the joints. Our approach avoids prior geometry, and prior manual processing of data. Our dynamics (Figure 9) place relatively few *a priori* spatial-temporal constraints on the movement, so the textures can be moved autonomously by the participant, and with more painterly gesture.

Jacquemin [14] used a technique by Georgii and Westermann to implement particle dynamics as texture computations on the GPU, in order to animate graphs. Our strategy is to skip the translation into and out of the relatively sparse graph representations required by particle models, and work directly at the full density of a bitmap.

Although it is natural to use increased computational resources to simulate 3D graphics and 3D physics (see for example Irving et al. [13]), we focus on 2D techniques to maximize machine resources and algorithmic power, but also for the key reason that we are delivering our results through two-dimensional projection devices: displays and video projectors. In our experience working with participants in live responsive environments, the experiential, i.e. phenomenological, impact of causal models of three-dimensional physics is dwarfed by the impact of kinesthetically plausible, perceptually negligible-latency dynamics on two-dimensional surfaces.

In 3.3, we detailed how our computational physics relates to prior work by Stam and other authors (including [23], [15], [11], [5]).

5.2 Physical Models Leverage Body Intuition

Replacing cognitive models to be mastered by the user with simulations that have shallow semantics but rich physics to produce textures that can be driven at video rates in response to user movement and gesture. Such responsive video textures can be projected as fields of structured light that can be shaped as palpable matter by the inhabitants of the environment.

The heat and Navier-Stokes models offer rich affordances, *continuous in space and time*, by which the user can manipulate video as structured light. These leverage his/her own tacit, lifelong, corporeal experience of whole classes of physical material like water or smoke. However, the Ising model for magnetic domain does not

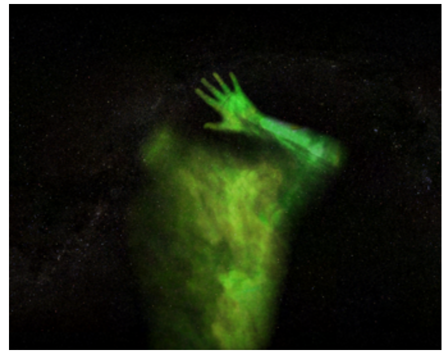

Figure 12: Movement artist working under elastic membrane illuminated by realtime video synthesized from his movement. We blend physical, topographic shadow with projected dynamical texture.

seem as accessible, because it references magnetic material, which is familiar to most people in the form of compact magnets, rather than continuous volumes of distributed magnetic matter. It is a subtle question whether and how a model such as the Ginzburg-Landau equation, whose primitive – a spin vector – corresponds to nothing perceivable by human senses, could scaffold intuitive manipulation at the macro-scopic scale of the phenomena it simulates, in this case a piece of generalized magnetic material. In any case, what we gain by using computational simulation rather than the actual physical material itself as the interface medium, is the potential for modifying the behavior of the computational matter to non-ordinary, *quasi*-physical behavior for expressive purposes, while still accessing the participant's physical intuition. We can then design new forms of matter that respond in interesting, and dynamically varying ways to participant activity.

5.3 Future Work

We are exploring classes of movement and gesture to condition parameter spaces and mappings that yield powerful, interesting, and evocative computational material. We are also exploring higher order phenomena. For example, adding temperature and quenching temperature as parameters can offer a sense of *phase change*, a class of material phenomena that we are exploiting more systematically in our Ozone media choreography framework.

We are planning more elaborate, professional applications of some of these calligraphic video techniques working with choreographers and movement artists on one hand, and with architects on the other. (Figure 13) One is an augmented performance work with the Blue Rider contemporary music ensemble [1]. And another is an experiment on memory and the built environment.

There are several promising directions in which to expand on our computational approach. We are interested in improving the simulation by working with velocities located on the sides of the grid cells [8], using a level set formulation to simulate the interaction between fluids and air [5], taking advantage of data level parallelism as found in modern processors, and distributing the simulation across a network of machines to provide richer detail.

Notes and Comments.

Michael Fortin was responsible for the implementations of the Navier-Stokes model, Yannick Assogba for the Ising model, and Jean-Sebastien Rousseau for the design and implementation of the video processing framework. We thank the other members of the media choreography team: Tim Sutton, Harry Smoak, Morgan Suther-

Figure 13: Exterior LED lighting parameterized via DMX by activity using live video and sound processing. TML installation at the Canadian Centre for Architecture, Montréal, 2009.

land, and Emmanuel Thivierge, and the TML [20]. We thank Prof. Peter Grogono, the Computer Science Department, the Faculty of Fine Arts, and Hexagram for their support. And we thank Sooyeon Cho for working on movement and choreography related to this research.

6. REFERENCES

[1] L. Adams, M. Montanaro, and X. W. Sha, “Frankenstein’s Ghosts (Blue Riders and Topological Media Lab).”

[2] K. M. Barad, *Meeting the universe halfway: quantum physics and the entanglement of matter and meaning*. Durham: Duke University Press, 2007. [Online]. Available: http://www.loc.gov/catdir/toc/ecip0619/2006027826.html

[3] A. J. Chorin and J. E. Marsden, *A Mathematical Introduction to Fluid Mechanics*, 3rd ed., ser. Texts in Applied Mathematics. Springer, 1998.

[4] J. Cohen, M. Olano, and D. Manocha, “Appearance-preserving simplification,” in *SIGGRAPH ’98: Proceedings of the 25th annual conference on Computer graphics and interactive techniques*. New York, NY, USA: ACM, 1998, pp. 115–122. [Online]. Available: http://www.cs.unc.edu/ geom/APS/APS.pdf

[5] K. Crane, I. Llamas, and S. Tariq, *GPU Gems 3*, 1st ed. Addison-Wesley, 2007, ch. Real-Time Simulation and Rendering of 3D Fluids, pp. 633–675.

[6] *jit.repos Reference*, Cycling ’74. [Online]. Available: http://www.cycling74.com/docs/max5/refpages/jit-ref/jit.repos.html

[7] P. Dourish, *Where the action is: the foundations of embodied interaction*. Cambridge, Mass.: MIT Press, 2001. [Online]. Available: http://www.loc.gov/catdir/toc/fy036/2001030443.html

[8] R. Fedkiw, J. Stam, and H. W. Jensen, “Visual simulation of smoke,” in *Proceedings of ACM SIGGRAPH 2001*, ser. Computer Graphics Proceedings, Annual Conference Series, Aug. 2001, pp. 15–22.

[9] E. T. Gendlin, *Experiencing and the Creation of Meaning: A Philosophical and Psychological Approach to the Subjective*. Northwestern University Press, 1997.

[10] R. Gerber, A. J. Bik, K. B. Smith, and X. Tian, *The Software Optimization Cookbook*, 2nd ed., D. B. Spencer, Ed. Intel Press, 2006.

[11] M. J. Harris, *GPU Gems - Programming Techniques, Tips, and Tricks for Real-Time Graphics*. Addison-Wesley, 2004, ch. Fast Fluid Dynamics Simulation on the GPU. [Online]. Available: http://http.developer.nvidia.com/GPUGems/gpugems_ch38.html

[12] E. Husserl, *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy*, ser. Collected works / Edmund Husserl. The Hague: M. Nijhoff, 1982, vol. 1<2>.

[13] G. Irving, E. Guendelman, F. Losasso, and R. Fedkiw, “Efficient simulation of large bodies of water by coupling two and three dimensional techniques,” *ACM Trans. Graph.*, vol. 25, no. 3, pp. 805–811, 2006.

[14] C. Jacquemin, “Allegra: a new instrument for bringing interactive graphics to life,” in *MM ’08: Proceeding of the 16th ACM international conference on Multimedia*. New York, NY, USA: ACM, 2008, pp. 961–964.

[15] T. Kim, “Hardware-aware analysis and optimization of stable fluids,” in *I3D ’08: Proceedings of the 2008 symposium on Interactive 3D graphics and games*. New York, NY, USA: ACM, 2008, pp. 99–106.

[16] L. Loke, A. T. Larssen, T. Robertson, and J. Edwards, “Understanding movement for interaction design: frameworks and approaches,” *Personal Ubiquitous Comput.*, vol. 11, no. 8, pp. 691–701, 2007.

[17] H. R. Maturana and F. J. Varela, *Autopoiesis and Cognition : The Realization of the Living, Boston Studies in the Philosophy of Science, Vol. 42*. D. Reidel Pub. Co., 1980.

[18] F. Neyret and N. Praizelin, “Phenomenological simulation of brooks,” in *Proceedings of the Eurographic workshop on Computer animation and simulation*. New York, NY, USA: Springer-Verlag New York, Inc., 2001, pp. 53–64.

[19] J. Petitot, *Morphogenesis of meaning*. Bern: P. Lang, 2004, vol. 3.

[20] X. W. Sha, “Topological media lab,” Concordia University (Georgia Institute of Technology 2001-2005), Tech. Rep., 2001–2006.

[21] ——, “Resistance is fertile: Gesture and agency in the field of responsive media,” *Configurations*, vol. 10, no. 3, pp. 439–472, 2002.

[22] X. W. Sha, T. Sutton, J.-S. Rousseau, H. C. Smoak, M. Fortin, and M. Sutherland, “Ozone: A continuous state-based approach to media choreography,” Concordia University, Topological Media Lab, Tech. Rep., 2009, submitted to ACM Multimedia 2009.

[23] J. Stam, “Stable fluids,” in *SIGGRAPH ‘99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques*. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1999, pp. 121–128. [Online]. Available: http://www.dgp.toronto.edu/people/stam/reality/Research/pdf/ns.pdf

[24] R. Thom, *Mathematical models of morphogenesis*. Chichester: Ellis Horwood, 1983.

[25] ——, *Structural stability and morphogenesis: an outline of a general theory of models*. Reading, Mass.: Addison-Wesley Pub., 1989.

[26] J. Varona, A. J. i Capó, J. Gonzàlez, and F. J. Perales, “Toward natural interaction through visual recognition of body gestures in real-time,” *Interacting with Computers*, vol. 21, no. 1-2, Special issue: Enactive Interfaces, pp. 3–10, January 2009.

[27] Y. Visell, “Tactile sensory substitution: Models for enaction in hci,” *Interact. Comput.*, vol. 21, no. 1-2, pp. 38–53, 2009.

[28] A. Webb, A., E. Koh, P. Joshi, Y. Park, and R. Graeber, “Choreographic buttons: Promoting social interaction through human movement and clear affordances.” Santa Barbara: ACM Multimedia, October 2006.

[29] K. A. Woolford and C. Guedes, “Particulate matters: generating particle flows from human movement,” in *MULTIMEDIA ’07: Proceedings of the 15th international conference on Multimedia*. New York, NY, USA: ACM, 2007, pp. 691–696.

[30] C. Woolley, *GPU Gems 2 - Programming Techniques for High-Performance Graphics and General-Purpose Computation*. Addison Wesley, 2005, ch. GPU Program Optimization. [Online]. Available: http://http.developer.nvidia.com/GPUGems2/gpugems2_chapter35.html