

Improvisational Media Fabric: take one

Keynote address to APDC

Lisbon, Portugal

September 26, 2002

Glorianna Davenport

Principal Research Associate

Director, Interactive Cinema

MIT Media Lab

Co-founder, Media Lab Europe

gid@media.mit.edu

<http://ic.media.mit.edu/>

"The structure of expression is shaped at the intersection of technology and culture"

Story-making is an activity of intelligent guessing, an expressive exploration of constructed meaning. Storytellers transform real-life observations into narratives through acts of selective inclusion, synthetic emphasis, time distortion, and metaphoric encapsulation; their tales are shaped within the empowering and framing constraints of their chosen medium. In art, the act of seeing and expressing outer and inner worlds becomes a shareable learning experience that is simultaneously culture-driven, culture-forming and culture-reflective.

Today I have chosen would like to share my ideas on the how the media will transform itself in the future into an improvisational fabric. Over the past two decades, the tools of cinema have become mass-marketed consumer goods and a substantial portion of the population is now empowered to record, reshape, and share aspects of life that interest them using an ever-growing inventory of palm sized video cameras and computers equipped with nonlinear editing systems. However, despite the growth of the World Wide Web, the expressive channels available for distribution of complex compilations of media material will depend on the integration of the full range of distributed networks.

In the past few years, this integrated communications channel has begun to emerge from the intertwining of networked computers, audiovisual displays, cameras, microphones, and activity sensors in the digital domain. In traditional Cinema, the activity required to make a film is radically different from that required to watch one. The cinematographer moves her camera through space and time so that the audience can remain in their seats, eyes fixed on an immobile view-screen. A reel of film spools through the projector from end to end, oblivious and unresponsive to the reactions of its audience. Today, digital cinema is freeing itself from its linear celluloid base; it is evolving into a "meta-cinema" where one's own memories, perceptions, actions, and desires connect with others through a continuous process of communicating, interweaving, and reconfiguring tradeable bits within a universal media environment.

Meta-cinema explodes the myth of the fixed and immutable heroic tale projected onto a single screen for passive consumption by a seated audience. New story forms become possible through its intrinsically programmable, re-configurable, and interactively responsive nature. The raw materials for a meta-cinema experience are likely to be scattered throughout the vastness of the Internet, waiting for a story-building engine (or

a person) to discover, select, arrange, display and trade them. While members of the audience may be widely scattered in a physical sense, they are networked together, connecting with story and with each other through shared channels. In meta-cinema, the audience – individually or collectively – can go to a story or the story can come to the members, at any time, wherever they are, whether stationary or traveling.

As meta-cinema abandons the single fixed screen to project itself into more and more of our everyday environment, it is becoming an improvisational learning partner that engages us in factual discovery, evolving hypotheses, augmented dialogs, and sociable interchange. In such an environment the differences among making, consuming, and sharing stories fade to insignificance.

Research and Future Cinema

Cinema has two parents: technology and art. In the commercial and policy vernacular of today's broadcast media dialog, Digital Television often references the current allocation of digital spectrum for high-definition broadcast television. As cell phones and other hand held devices with multimedia capability become available their social function will need to merge with programmable content and the "More button" in shaping the new cinematic story forms?

Using a broad and somewhat ideal definition, we can say that Interactive TV presumes that video/audio content has been created for transmission over a 2-way (broadband) digital channel, and that the form of its content assumes that the end receiver or audience will use the two-way capability of the channel for activities related to the program content. More simply, I suggest that iDTV assumes the integration of TV, computation, with the Net and mobile networks.

TV, broadly speaking, refers to video content that uses broadcast, cable or satellite technology to reach a broad base of consumers. Alternatively, we think of TV as a rather large device that sits in a prominent position in many living rooms around the world.

Computers are machines we can program to carry out particular procedures. When we combine television programs with computers we can process digital images and we can simultaneously use a 2-way communication channel to call up other information or to send messages. With certain computing devices such as TiVO, we can store incoming video content thus allowing us to time-shift program content to avoid advertising or to allow us to watch several simultaneously broadcast programs according to our own schedule.

The Net is an emerging distributed channel that has no central or hierarchical architecture. The net is a self-organizing network formed as signals from computationally enabled devices in geographic proximity find and use each other to pass messages and share power. The net is a mix of commercial and a large non-commercial carrier owned by consumers at various levels who own the required devices including servers and who choose to share or make these devices available to devices owned by other people. The net provides us a very new way of sharing media whose precedents we have already seen in architectures such as Napster.

The new media, in whatever form it is realized, will capitalize on the convergence of these three technologies: TV, computation, the net because they are available and more importantly because they all play a critical role in today's culture of media consumption.

Interactive TV: The movie I want, when I want, where I want ... ?

Surfing channels of TV is now a commonly acknowledged phenomenon. TIVO and other systems bring us closer to fulfilling the the consumer's desire to allow their personal life to interrupt and adjust their TV viewing. The "MORE" button provides the second most ubiquitous concept in interactive programming. The more button is fairly simple but tiresome to implement by hand coding, but, because MORE is a contextually dependent desire, it is less easy to program automatically.

The simplest case of MORE assumes that there is only one choice available and that this choice has been pre-scripted to be visible to the consumer either by means of active selection from a graphical menu or a hyperlink. This level of interaction can work well for Interactive Television systems today which have only a very narrow band upstream capability. "Interactive Julia", a collaboration between Boston's WGBH-TV and Michael Bove (plus his student collaborators, Stefan Agamanolis and Jonathan Dakss) provides a useful example of content recycling using a MORE button. The show opens as Julia Child, a legendary chef and TV personality, sweeps grandly a room of guests announcing: "Dinner is served." At this point the curious viewer might wonder: "What's on the Menu?" MORE takes the light-pen-wielding interactor to the Menu. From here, the interactor want to know about the entrée; MORE takes the interactor to Julia's discussion of how to select a filet of beef. Minus the more button, the interactor goes back to the main program content.

The MORE button found its extreme commercial example in "Hypersoap," a melodramatic soap opera which substitutes clickable on-screen product placement for advertising breaks. By pointing and clicking on any object in the scene -- a chair, a vase, the actor's sweater -- the viewer could call up a catalog page and ordering information for that item. After the video material was shot, it was extensively post-processed with special software which tracked, segmented, and hyperlinked each clickable object. The resulting interface was an invisible overlay of clickable target zones -- an interesting alternative to overlaid text menus. The experience of creating these two programs inspired Jon Dakss to found a company, Watch Point Media, whose mission is to provide a visual software authoring package that ill allow media editors to easily program simple instances of the MORE button.

Can a MORE button engage us in a deeper level of story interaction? Will this approach to content programming give rise to stories that really change our knowledge about ourselves and our world? Will it allow us to learn as we view? Will it allow us to participate in the creation as well as in the consumption of narrative?

From on-screen navigation to an out-in-the-world experience

The age-old story of the journey must inevitably lead to a world that knows about its spatial self. At the Media Lab and commercially, the earliest interactive spatial television experiment Aspen, the earliest spatial interactive television experiment has been transformed into the navigational systems of automobiles. We now ask how spatial sensing technologies such as GPS, field-sensing, wireless triangulation and IR

communications can be used to developed mobile, location-based stories. As we do so, we need to recognize that while it is difficult to predict the mental trajectory of a traveler, their path through space is continuous.

In contrast, the "old and the new" story model requires a different type of continuity (and choice). Stories about people highlight "conflict and resolution" models; order is not mapped to geography but to the mind, the emotions, and idiosyncratic memories of personal experience. How do we represent to the machine the underlying phenomenon that drives this story type -- the old ever resists the new and the new wants to clobber the old?

What interface characteristics allow us, the audience, to position our selves at the heart of the story world and make our way forward? What is "forward" in a story space? We cannot navigate toward our desired destination when we do not know that destination; we can only pick and choose from options that present themselves. If the audience lacks a sense of context and direction -- an overview -- use of these options can become quite arbitrary and short-sighted.

Over the past 20 years, only a few options have been presented that allow us to navigate interactive story worlds or webs: on-screen navigation, physical navigation, polling, direct interaction with a character, and tangible interfaces.

Briefly, on-screen navigation allows the audience to choose a direction from a menu of options. In Aspen, we could choose to turn right or left or stop or go inside; in New Orleans we could choose from a palette of context-dependent icons that updated itself regularly. With on-screen navigation, the audience must make an active, thoughtful choice whether or not they know where the choice will lead.

Physical navigation through space provides a more subtle approach. Here we can embed segments of the story web in real physical places, such as streets and buildings or the individual displays of a museum. As we move through a place, our proximity to these embedded information artifacts allows them be viewed immediately or collected for later use. Sometimes, our motion through a room is driven by an express desire to engage with its attractions and affordances; sometimes, it is an incidental by-product as we carry out the tasks of everyday life.

Polling or voting provides a more sociable interaction: we are part of a community; we are asked; we answer; we experience according to the opinion of the larger group. Typically, polling requires some discussion, some use of a communications network. This approach gains from the collective knowledge and experience of a broader audience, but suffers from the dilution of individual desire.

Tangible interfaces have long been a staple of interactive machine control. The panels of knobs and switches on the front of radios and TV sets have spawned a portable, hand-held doppelganger: the wireless Universal Remote. Many race-car video games offer plastic steering wheels and foot pedals -- which mimic the function of their real-world counterparts -- as physical, literal interfaces to the simulated game world. Tangible interfaces can also be highly figurative and poetic; for example, when the cork is removed from one of Hiroshi Ishii's first beautiful glass bottle the local Internet traffic at that moment plays out as pleasing sounds. Later one bottle became a three, and a new tangible interface to narrative was born in the "Genie Bottle" story.



The way we move through a real physical space can also be used as an input to story construction and playout. Installations in public spaces can give story a sociable characteristic; so can stories that invite viewers to communicate and share materials over the Net. This approach to story navigation can take advantage of a wide variety of sensing devices -- such as pressure-sensitive floormats, mini-radars, sonars, field sensors, vision systems, and GPS -- which are used to detect whether an audience is present in a space, determine their position, and report specific aspects of what the audience is doing

In the early 1990s, we began to explore how a full-body navigational paradigm could be used to extend story interaction and make navigation of the story space more transparent. In Sammy Spitzer's "Pigeons," an array of tiny sonar sensors detects whether someone has entered the space; if they have, the pigeons are "startled" and fly away. The most remarkable thing about this first, simplest experiment is that everyone who saw it laughed. This led us to wonder whether we could create more complex short stories for public space that would generate emotional reactions in the audience.

Reweavable media fabrics

"Improvisation is a creative activity where no boundary between cognitive and intuitive is drawn, where no clear rule-based approach can be utilized in even a superficial way, and where the focus is on the process rather than on some final goal or state."

Paul Nemirovsky, PhD Interactive Cinema [5]
<http://web.media.mit.edu/~pauln/research/goals.html>

As cinema frees itself from the limitations of the celluloid strip and the continuous viewing experience, meta-cinema emerges. For the past several years, researchers in Interactive Cinema have tussled in various ways with the vision that in the future, we -- as authors -- will generate an evolving collection of cinematic elements, that the interface will grow increasingly transparent and then disappear, and that a dynamic, re-

configurable, ever-present network will connect people with people and with media stories at whatever density, localized specificity, and temporal duration is appropriate to the context and inclination of the receiver. In this future cinema, we do not "go" to the movies; rather we are always in the midst of improvisational meta-movie potential and learning stories.

Barbara Barry's *Story Beads* (2000) presented a clear alternative view to shared media in the future: its central mechanism was a wearable necklace of media beads. Each bead was a wireless transmitter/receiver specialized for image transfer and storage. If you wanted to review or trade your pictures, you could send them to the central video amulet; in a trade, you simply held up your amulet to another and the picture would jump across. The idea grew out of discussions with young girls who wanted to maintain some control over who was able access their images. [6]

In *Tangible Viewpoints*, (2000-present) Ali Mazalek explores how physical objects and augmented surfaces can be used as tangible embodiments of different character perspectives in a multiple viewpoint interactive tale. By providing a direct, tangible mode of navigation to the story world, we seek to bridge the gap between cyberspace and our physical environment as we engage with digital stories. The table upon which game-pieces can be moved is also a projection surface which invites several people to gather around the story space to input their own story or otherwise to browse through a pre-formatted story. Recent extensions to this environment include the use of surround sound and graphical spatial maps. [7]

Throughout these experiments, the relation between sociability and technological progress remained murky. Inspired by thinking around an earlier embedded work, *Pigeons* (1996) [8], *Flights of Fantasy* focused on the idea that a message is never the same for the sender/maker and the receiver. This project was a two-room installation in the DeCordova Museum during the Boston Cyberarts Festival in the spring of 2001. The visitor began their interaction in whichever room they first entered. In one room, the audience approached an ornate wooden table whose surface resembled a giant version of a child's pocket puzzle. By sliding the symbolized blocks of this interface, the visitor created a path of icons; an action-selection program translated that path into a sequence of video segments and text narration drawn from a database. The table interface alleviated much of the self-consciousness that might otherwise accompany public creation of a message. At the opening, one woman commented to her friend, "Look, I made that movie!" The table masked the complexity of the editing activity, making it informal, playful, and message-like. [9]

In an adjacent room, the visitor entered a forest of birdcages, each housing one or more miniature video screens. In an idle idyll state, a video of birds looped on the small screen. When someone opened the door to one of the cages, the bird video disappeared and was replaced by one of the messages created by a visitor in the other room. As these messages flew from room to room, they were accompanied by the sound of flapping wings. One authored clip was a movie of a woman tearing a photograph into bits; as she throws it out the window, the visitor hears an audio clip selected by the system logic, such as: "Would you trade a piece of piece of plain drawing paper for a book of poems? Really?" The somehow familiar video stirs an ad hoc memory in both the maker and the receiver, although it is only at the receiving side is the video linked to an evocative audio sample. *Flights of Fantasy* builds upon earlier works in which we explored the idea of responsive media in casual architectural space. Again, the experience of sending and receiving is held purposefully distinct.

Most recently, we have taken the idea of time/space discontinuity to heart in the implementation of M-Views, a contextually-aware ad-hoc platform for mobile cinema. Equipped with an IPAQ PDA, we wander through MIT's campus allowing cinematic sequences to come to us. These sequences are a mix of fictional, documentary, historical and futuristic story accounts of life on campus. For instance, in "Another Alice" the GPS tells the system where the consumer is and the story time is mapped to actual real-world time. As the consumer meets characters, the characters tell her where she might want to go next. The fun of "Another Alice" has something to do with how people who have experienced the story relate their experience to others.



Conclusion

The "old culture" of TV was based on the notion that channels are a scarce resource. Pre-made chunks of content -- TV programs -- were distributed to a mass audience via a shared "front-channel." Content flowed according to the schedules of "broadcasters." Any interruption or discontinuity in this monolithic feed was considered a disaster. The audience's control was limited to turning their receivers on or off and selecting which channel to play. A vast divide was created between content makers and consumers, with no provision for direct, real-time audience feedback.

The "new culture" of interactive TV will be based on the notion that virtual channels are an abundant resource, that content can come from anywhere in the highly-distributed Net, and that a "back-channel" of related communications (whose traffic may dwarf that of the original TV program) will be intimately associated with the primary "front-channel." Increased switching and routing capabilities will allow programs to be fed out on a shot-by-shot basis, allowing a degree of personalization and customization hitherto unseen. Object-based channels can be streamed and composited into a single scene, greatly expanding the reusability and repurposing of the underlying databases of content.

Throughout the history of art, literature, film, and television, whenever a compelling chunk of content (or an episodic series of them) was published, communities and subcultures gathered around it. The peripheral activities of content providers (such as "behind the scenes" stories and scheduling information) and the activities of "fans" (such as gossip, speculation, and the exchange of relics) have increasingly moved on-line; soon, they may be integrated with the primary channel itself. The limits of community participation may be economic rather than cultural or technological: who will pay for the back-channel? How will the commercial framework constrain access and story form?

In developing new media forms, we need to look at how computation and communication is changing culture and grow an understanding of the true nature of the Net. We also need to ask what stories need telling, and by whom; what story structures are inherently suited to computational media; how computational narrative resonates with the technological progress; and what, if anything, technological fluency means for a marketable product.

As audiences gain substantial experience with interactive interfaces, their learning curve will decrease and their use of the system will become more sophisticated. Their relationship with content will shift from today's relatively passive enjoyment to highly active, idiosyncratic, and context-aware encounters. To facilitate their mastery of control, the audience should be provided with some sort of contextual overview of the available materials; as we learned with the New Orleans project, audiences were happier and more discriminating when first presented with a short version of the story before they dove into the full depth and breadth of the available content. As the primary content-makers, audience, and mediating machinery work together in the co-construction of meaning and shared experience, limited look-ahead will become a crucial mechanism of content navigation and ployout.

The era of the Set-Top Box sitting atop a huge, immobile TV set in one's living room will quickly pass as mobile devices become more powerful, ubiquitous, and interconnected with each other and their environment. Then, the excitement will come from creating beloved forms of tangible interfaces for hyperlinking, public screen spaces for planned and accidental encounters, and mobile outreach for new story forms.

When centralized and local computing power is merged across high-speed networks, program content will truly become an "open ended invitation to ethical and poetic responsiveness"