# NEXT GENERATION INTERFACE FOR MULTIMEDIA PUBLICATIONS

Glorianna Davenport
Assistant Professor of Media Technology
The Media Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts, USA

Prepared for: '91 Friend21 Symposium

## Abstract

Future multimedia publications will consist of content, identified as the stuff which invites interpretation. and interface. identified as all computational elements which orchestrate delivery. This paper focuses on information architectures for electronic multimedia publications which include digital video as a principal carrier of content. These publications will rely on multiplexed secondary databases to augment content. Images will be composited and sequenced at run-time in the delivery environment. User interactions will be conversational and will enable annotation to personal data structures.

## 1.0: Introduction

As individuals we all use interaction and conversation as basic methodologies for learning. Over the past century, the role of visual and audio media in communication has grown and in some domains now competes with text. Today there are many topics of concern which could be better understood if video. sound. text and computational elements were bundled such that the subscriber experienced a flexible and personalized presentation. Many publishers have begun to acknowledge the importance of electronic multimedia forms and are trying to understand potential opportunities for early entry into this market.

How should electronic publications be structured to assure delivery of quality information to the subscriber as well as economic profit for both publisher and author? Information gathering. media production. interface design and programming are expensive endeavors. As with traditional journals. a template needs to be invented which streamlines the structuring of multimedia publications and their production methodology.

Over the past decade. academics have prototyped a variety of computational multimedia courseware products [1] in a range of disciplines including science, history, architecture. an, literature, language learning, business and international affairs. While many of these experiments have been enthusiastically reviewed. they have for the most pan been one of a kind publications. The content of these products has not been directed at the interests of a broad subscriber base. nor has the production methodology been subject to the demands of serialization. publication deadlines or commerce. However there has been sufficient

experimentation to suggest some guidelines for new interfaces and suggest a more streamlined methodology which will allow video document collection to be annotated in the field and to be more quickly integrated into an electronic environment.

## 2.0: The ABC's of Multimedia Publishing

In order for multimedia publishing ventures to become economically feasible, four requirements must be met. First and perhaps most urgently, a product must be developed that will be purchased by a critical mass of people. Second, a template must be invented which invites subscriber participation and streamlines production. Third, a common and recognized format for multimedia signals must be agreed to so that the same material can be displayed on multiple hardware platforms. In image compression, this "standardization" has all but been achieved; however, standards for content description may be more difficult to achieve. Last, extensible tools which enhance the process of creating multimedia elements must be developed.

What is this new medium? How might we want to interact with it? Pervasive, fun, shareable and interactive are some characteristic qualities. Similar to TV in that I turn it on when I come home, but unlike TV in that I can affect what I choose to see. For instance I can browse and sample movies at will; in addition to viewing excerpts, I can enjoy the comments of my favorite movie reviewer. I can select a specific journalist to create a perspective on today's news. I can browse product selections -fabrics, wallpaper, clothes, books, food and even medical supplies. Finally, should I want to relax I might ask to participate in a narrative in which my profile is used to affect action from the outset.

What remains to be designed is a well crafted "form follows function" conversational interface which supports reliable limited look ahead functions as well as intuitive tools for annotating and personalizing content. This is not a solved problem. Descriptions of on-going or recent research will be used to illustrate aspects of what may be incorporated in the next generation interface.

## 3.0: The publishing landscape: issues for the principle players

As we have mentioned, the success of such electronic publications will depend on the economics of information generation and distribution. Today the situation for multimedia publication is rather different than it was during the 1980's when the bottleneck to invention was linked to hardware capability and cost. Today, almost any technical specification can be met; rather than arguing about which technology for how much, we need to ask what products at what cost and how appealing they will be?

### 3.1: Issues for the publisher

Before publishers can commit to developing and distributing- multimedia publications, they must understand what they are producing. Publishers must feel confident about the

look and feel of the publication. They must know what they are selling as their primary product. In the electronic age, publishers will struggle at both the interface and the content level with a buy or make problem. At the interface level, this decision will impact the scale and profitability of distribution. At the content level, many publishers may be able to take advantage of peripheral products when they are deemed sufficiently valuable as a means of lowering overall production cost. This idea needs to be further explored. For instance, publications such as community news suggest value will be added if we can automatically link between primary source material (features) and secondary source material (such as map sets and yellow pages.)

### 3.2: Issues for the content producer/author

Unlike printed texts where all information is packaged on a set of consecutive pages, computational environments suggest that material must be conveniently granular such that elements can be added or omitted and the general order can be re sequenced as necessary. Modular construction requires explicit understanding of the linking mechanism, knowledge of the secondary material, and skill in creating content chunks which can be assembled in various ways. Generating a quick turn-around on stories will require a team effort between individual content producers, managing editors, interactive designers and engineering staff. "Thick" [2] content descriptions and availability of story templates will enrich both the author's craft and the machines ability to generate a personalized scenario for consumers.

### 3.3: Issues for the subscriber

For the subscriber, the quality of both content and interactive experience are critical. In the simultaneity of the delivery environment, the two should in fact be indistinguishable. As the consumer gains familiarity with the machine, the machine must distinguish the computer-literate trail blazer from a johnny-come-lately to the technology, a story expert from a story novice. In addition, the machine must learn how to distinguish user input cues which could indicate boredom from those which indicate confusion, and be able to vary the style, pacing, depth or range of content based on limited user dialog and knowledge about available source information.

## 4.0: The Community News Model

Community news provides us with a provocative thought experiment. Community news supports a semi-structured information set with content subclasses which are recognized by their familiar form and explicit function: features, listings, advertisements, maps, photographs. Transcoded into an electronic counterpart, it is easy to imagine how content and functionality can be scaled up or down depending on subscription rates or capability of the hardware platform; we can also imagine how content can be shared across content subclasses.

The vision which is required to produce an electronic "infotainment" publication is not unlike the marriage of talents which generated Life magazine. The contributions of

multiple industry players to such an undertaking will break new economic ground in the publishing arena. New learning interactions based on games and puzzles will develop naturally give the nature of electronic interaction- and herald a generation of new content design and user/machine dialog.

### 4.1: Multiplexed Databases

How can we make the creation and delivery of multimedia experiences economically viable in the coming years. Like syndication of the past three decades, a variety of motion picture, sound and graphic databases could be generated by specialized publishers who are practiced in evolving content for these domains. Examples include historic monuments, video and text news, real estate databases, yellow page equivalent, scientific glossaries etc. By providing standard hooks in the indexing of these publications, publishers could consider news a secondary market for their products; by generating automatic links between the first level and second level materials, the consumer would have available a much denser information set and would pay on a per use basis for that expanded information.

This method of aggregating content has certain economic consequences: 1) it can amortize the cost of generating and disseminating information by creating secondary revenue sources; 2) it supports anew genre of literary artist who can benefit economically by infusing the subject matter with a particular point of view, namely his/her own expert perspective. In the end, the consumer benefits from the ability to explore what they want, when they want. Assuming reasonable availability of expert structures, smart limited look ahead functions, and easy to learn conversational interfaces, these secondary databases offer publishers a major win in developing a diverse market.

### 4.2 Layered Advertising: a special content case

In evaluating the benefits of multiplexed databases as an economic structure for electronic publishing, the design of advertising must be considered as a special case. A key question for future systems is who will pay for programming and distribution of content. The TV model hooks advertisers, while the cable model hooks subscribers. This author suggests that all parties, including the consumer, the distributor, the publisher and the advertiser, will pay something. In the end the consumer will only pay when s/he values the information received.

Clearly unsolicited information, which can be considered by far the largest class of advertising today, can be beneficial. However often it is unwanted and wasteful. In an electronic publication, zapping advertising will be a common practice. In cases where advertising provides an information service which consumers want and need, they will pay indirect proportion to their interest. Layered advertising encourages a multilayered structuring of message by advertisers. Short visual commercials can provide visibility, while deeper informative messages will be considered a service. In this sense, layered advertising supports both the provider and the extractor of information within a formula that is feduciarily viable.

Consider the yellow pages as it will exist electronically tomorrow. We imagine that the phone company will continue to feel that it is in their best interests to publish a directory. Today, if you search on bicycles, you will discover Fred's Bike Shop, with address and phone. In the yellow page listing, it is surrounded by similar stores. In the electronic version of the yellow pages, Fred's Bike Shop can add a 15 sec. commercial to the listing for a price. The advertisement shows up as a small, postage stamp window and carries with it sound, music, lingo, text, and motion picture. The consumer does not pay extra to view this. However they may pay when they probe further to receive the daily tip or a bike tour of their choice.

Similar to yellow pages today, the phone company picks up the tab for the listing, the advertiser picks up the tab for the video commercial, the bike shop and, consumer split the cost of the "tips" and the consumer picks up the entire cost of the "bike tours" which is in itself an integration of footage provided by the bike shop and landmarks which are pulled from a database of historic landmarks which mark the selected route. The key here is, any bike tour can be made up on the fly specifically for the customer.

### 4.3: Journalistic Methodology and Creative Tools

Both news gatherers and an increasingly broad spectrum of scientists use video on its own or with text annotations as a research tool. Today these scientists are forced to pick a single medium in order to tell their story. Electronic multimedia will allow scientists to gain a far more direct experience by combining video, text and graphics with interaction. However, this requires a generation of new tools which will insure that content producers can work effectively across media. [3]

Field annotation, multi-stream displays, and previewing devices some of the critical tools for multimedia production. For instance scientists tend to collect many hours of observations. A scientist's notebook would be enhanced if the scientists could annotate the video stream while he/she is recording picture and sound; annotation in partnership with soft templates or story models should generate meaningful image sequences to novice or professional observers.

The computational base in multimedia environments must be able to select shots and transitions, layout sequences and position user cues. These actions may need to be more or less automatically orchestrated, depending on program form, content and audient. Interactions of all sorts require some level of tools; we have tried to focus on tools which benefit makers and consumers equally. We do not consider these tools editors but rather construction makers. In order to generate the best possible selection of content for a particular users, these tools must encourage and, at some level interpret, user input; this input must then be reflected back so as to modify the plan or story model during the run time cycle.

Annotation and representation of content to the system must be accomplished before we can apply principles of interactive cinematics to a story building module. Representation is a difficult problem and we have chosen to approach it from multiple perspectives.

Image [4] and sound [5] processing modules can act as effective mediators in segmenting and filtering original materials.

Shared scientific databases where some material may be generated in the field, and were material frequently must be annotated to support collaborative, long distance observations is a growing application area. "Stratification" [6] is a method which allows us to annotate the video stream quickly using a system of keyword classes and keywords. We feel that this methodology will be particularly
powerful in cases where large collections of scientific material are generated as they are being used by multiple researchers as well as for most general databases of motion picture imagery where there is no single designated use associated with the imagery.

In the case of narrower content domains, including a later stage of the scientific example detailed above, we feel that a knowledge based approach can offer enormous advantages. The "Electronic Scrapbook", [7] was built to help amateurs edit their home movies. This system promotes inferences about major characters, places and events. The system is built on a representation language called Arlotje [8] which was developed by Professor Kenneth Hasse at the Media Laboratory. A new version of this language is currently in development and future work will involve linking Stratification to the Electronic Scrapbook interface as well as applying some of Professor Hasse's new work in episodic story understanding.

Despite this work, annotation is still a major stumbling block in creating well indexed multimedia collections. Both the author's experience in creating the "New Orleans in Transition" [9] and Ricki Goldman Segall's experience with "Learning Constellations" [10] underscore the importance of developing a methodology to ease the bottleneck associated with content description of temporal media. Any such methodology must be acceptable to the scientist. That is the annotation tools must allow rapid and accurate entry in the field or in the laboratory. The annotations must be able to be integrated with other available data, and must result in thick descriptions which support interesting scientific inquiry about the observations embedded in the content.

In addition to annotation for documentary, building a new generation of multi-threaded narratives is of increasing concern. Such narratives are possible and promise to be enormously entertaining. Game systems such as Nintendo suggest potential devices for interaction. However in order to build such complex narratives using dramatic conventions, we need tools which will allow us to previsiualize shots and track production of the complex scenarios. An early system is already beginning to affect how narratives are made. "The Slipstream" [11] allows makers to track on the set production with all available continuity data. Currently we are building a device we call the "Filmmaker's Eyeglass" which will allow directors to preview composited scenes on the set.

Much of our recent work in 3D computer graphics systems will be integrated into a 3D graphical storyboarding environment which can play out complex scenarios, showing required computational matches in the complex multithreaded narrative.

### 4.4: Conversational Interaction

In a real sense, the most interesting developments for multimedia narratives exist in the computational interactions between narrative agents and the participant-users. Our first explorations in creating participatory strategies were based on learning.

In an early urban planning case study, students introduced characters using video excerpts in their position papers. In some cases they were able to construct an abstract argument illustrating the immutable positions taken in a particular development argument. While this was not a strategy which enticed novice users to interact, these papers foreshadowed the next phase of designing enticing interactive strategies. Most consumers enjoy science features such as Nova or National Geographic. However the passive viewing environment of television does not allow most viewers to learn as much about underlying principles as they could or would like. Adapting Seymour Papert's view of "heads in and hands on" learning, moved us to consider the use of participatory computational puzzles or games as a natural extension for a science feature. [12] A rule base allowed us to strengthen student understanding of business strategy in a business case we developed in 1990. [13]

In our most recent experiments, we are exploring how the concept of microworld can be developed as a arena for interactive cinematics: for instance, by modifying the characteristics of a microworld, users can change the way in which that world and its inhabitants function. This sort of interaction is designed to create stronger relationships between the participant viewer and the narrative which is being constructed.

### 5.0: Conclusions:

We are still in the early stages of developing a richly textured information space which generates both meaningful associations through the use of spatial as well as temporal placement. Creating a conversational interface which is supported by cinematic juxtapositions is critical if we are to invite our subscribers to participate in shaping their own learning space.

Thought experiments such as Community News, including yellow pages and maps, allow us to understand the economic implications of using shared databases to people complex information spaces. Current research in annotation and story representations will allow systems to have knowledge of their contents and ultimately better limited-look-ahead functions. Microworlds present perhaps the best current model of how we might be able to orchestrate interaction and multimodal stories in an interesting and forceful manner.

**1**. Initial computational multimedia projects include "Aspen" a surrogate travel prototype, 1979-81 and "Automatic Transmission Manual" 1981-3, both developed at MIT's Architecture Machine Group; experiments in visual courseware for language learning,

urban planning, brain science and mechanical engineering followed, many of which were developed under the auspices of Project Athena at MIT in the mid and late 1980's.

**2.** see Ricki Goldman Segall, *Learning Constellations: a Multimedia Ethnographic Research Environment Using Video Technology for Exploring Children's Thinking*, MIT Ph.D. Dissertation, 1990 for a discussion of Clifford Geertz and Thick Descriptions.

**3.** Scientific case studies developed in the Interactive Cinema Group of the MIT Media Lab include Davenport. "New Orleans Interactive," 1983-91; Ricki Goldman-Segall, "Learning Constellations", 1988-90 Thomas G. Aguierre Smith, "Anthropologist's Notebook," 1991-92, prototypes which are designed to emphasize informal interactions and browsing include "Elastic Charles: a Hypermedia Journal," 1988; "SNAP: a Journal of Science and Technology," 1989, and "Elastic Boston" 1990-present.

**4.** Hero Ueda, Conversations and laboratory presentations, 1990-91.

**5.** Natalio Pincever, *If You Could See What I Hear,* MIT, MS thesis, 1991.

**6.** Davenport, Smith, Pincever, "Cinematic Primitives for Multimedia: Towards a more profound intersection of cinematic knowledge and computer science representation," IEEE Computer Graphics and Applications, Summer, 1991.

**7.** Amy S. Bruckman, *Electronic Scrapbook: Towards an Intelligent Home Movie Editor,* MIT MS thesis, 1991.

**8.** Ken Hasse, "Arlotje," 1990-1991.

**9.** Wendy MacKay and Glorianna Davenport, "Virtual Video Editing in Interactive Multimedia Applications," Communications of the ACM, vol. 32, no. 7, July 1989.

**10.** Segall, op cit.

**11.** Alan Lasky, "The Slipstream: A Data Rich Production Environment," MIT, MS Thesis, 1990.

**12.** "SNAP: An Elastic Journal of Science and Technology," "Weird Creatures Interface," Authors A. Bruckman and N. Farber, Info sheet, Interactive Cinema Group, MIT Media Lab.

**13.** Davenport, Harber, "Numbers: A Medium that Counts, IEEE Computer Graphics and Applications, July 1991.