HyperCafe:
Narrative and Aesthetic Properties of Hypervideo

Nitin "Nick" Sawhney[†], David Balcom[†] and Ian Smith[‡]

The Georgia Institute of Technology
School of Literature, Communication, and Culture[†]
College of Computing[‡]
Atlanta, GA 30332-0165 USA
{nitin, dbalcom, iansmith}@cc.gatech.edu

ABSTRACT

HyperCafe is an experimental hypermedia prototype, developed as an illustration of a general hypervideo system. This program places the user in a virtual cafe, composed primarily of digital video clips of actors involved in fictional conversations in the cafe; *HyperCafe* allows the user to follow different conversations, and offers dynamic opportunities of interaction via *temporal, spatio-temporal* and *textual* links to present alternative narratives. Textual elements are also present in the form of explanatory text, contradictory subtitles, and intruding narratives. Based on our work with *HyperCafe*, we discuss the components and a framework for hypervideo structures, along with the underlying aesthetic considerations.

KEYWORDS: Aesthetics, multi-threaded narratives, navigation, temporal links, digital video.

A VISIT TO HYPERCAFE

You enter the Cafe, and the voices surround you. Pick a table, make a choice, follow the voices. You're over their shoulders looking in, listening to what they say—you have to choose, or the story will go on without you.

Welcome to HyperCafe. The experience, the aesthetic, is choice: you decide when to listen, when to leave. You decide. The camera moves from one table to the next, and opportunities (what, in another context, J. Yellowlees Douglas would call a "narrative of possibilities" [11]) present themselves to you. Select a conversation, and the navigation pan fades into a close-up, two men talking, and one's saying to the other, "in fact, our words over here don't affect their words over there."

Another table comes into view, another possible conversation, next to the first. After a few moments, the second table fades away, if not selected. The link unrealized (which is a choice in itself), the story continues: "I find that highly questionable," the other man says. "What if I yell fire? What then? It seems my words have a great effect on [motions to table behind him] these people over here [motions to another table], these people over here."

Another table comes into view. A man with a thin beard is standing over a young woman with blond hair. You choose the second table and the first story fades back. Thin beard begins to speak. "Do you remember me?" he asks the blond woman. As possible narratives are realized by your touch, and the story forms around your choices, the question lingers. Do you remember me?

Another table, another opportunity to move.

INTRODUCTION

"Hypervideo" is digital video and hypertext, offering to its user and author the richness of multiple narratives, even multiple means of *structuring* narrative (or non-narrative), combining digital video with a polyvocal, linked text. We have redefined the notion of links for a video-centric medium, where they become spatial and temporal opportunities in video and text. In this paper we will discuss *HyperCafe* as the basis for a broader discussion of the narrative and aesthetic structures afforded by hypervideo. These structures and navigational methodologies later help us develop a conceptual framework for hypervideo. Our aim is to provide new thinking for a new mode of expression, much as George Landow did for hypertext: "Hypermedia, which changes the way texts exist and the way we read them, requires a new rhetoric and a new stylistics." [21] Now, too, does hypervideo.

Related Work

Our primary influence is the hypertextual framework found within *Storyspace* [30] where the relationships among the spatially organized writing spaces become

part of the content, permitting a duality of the writing space [4]. By constructing our script in *Storyspace* (see Figure 8), we were able to exploit the duality of structure and content to create a representation of linked "narrative video spaces." *Synthesis* [28] based on *Storyspace*, was used to index and navigate analog video content associated with text in writing spaces. In *Hyperspeech* [2], recorded audio interviews were segmented by topic and the related comments were linked. Like *HyperCafe*, the system focused on "conversational interactions"; one user of the system felt that she was "creating artificial conversations" between participants.

Video-to-video linking was earlier demonstrated in the hypermedia journal *Elastic Charles* [8], developed at the Interactive Cinema Group (MIT Media Lab), where Micons (miniaturized movie loops) would briefly appear to indicate video links. This prototype relied on analog video and laser disc technology requiring the use of two screens. Digital video today permits newer design and aesthetic solutions, such as considered in the design of the *Interactive Kon-Tiki Museum* [22]. Rhythmic and temporal aspects were stressed to achieve continuous integration in linking from video to text and video to video, by exchanging basic qualities between the media types. Time dependence was added to text and spatial simultaneity to video. Yet, unlike *HyperCafe*, moving text was not utilized and video links were represented by pictures of the video on static buttons. Temporal *opportunities* in *HyperCafe* permit only a temporal window for navigating links in video and text, as an intentional aesthetic. The nature of the video content in *HyperCafe* allows us to consider new conventions for indication of temporal and spatio-temporal opportunities in hypervideo. Opportunities exist as dynamic video previews and as links within the frame of moving video. Control over the video content permitted use of additional camera techniques to add continuity between video to video links via "navigational bridges."

Time-based, scenario-oriented hypermedia has been demonstrated in *VideoBook* [26][27]. Here multimedia content was specified within a nodal structure and timer driven links were automatically activated to present the content, based on the time attributes. Hardman et al. [15] utilize timing to explicitly state the source and destination contexts when links are followed. Synchronizing media elements is both time consuming and difficult to maintain. A system called *Firefly* [9] allowed authors to create hypermedia documents by manipulating temporal relationships among media elements at a high level rather than as timings. In *HyperCafe*, we deal with the presentation of temporal links within a continuity based on film aesthetic. We later discuss future work towards a toolkit for specifying such temporal links in hypervideo. These temporal links permit the possibility of presenting alternative (multi-threaded) narratives in hypervideo. *Agent Stories* [3] was developed at the Interactive Cinema Group (MIT Media

Lab) as a tool for creating multi-threaded story structures, built on knowledge representation of characters in the story. A multi-threaded documentary was demonstrated in *CyberBELT* [5], where the documentary evolved with feedback from the user, based on dynamic weights assigned to video clips. Although user-defined narratives are utilized in *HyperCafe*, our focus is on the "presentation" of aesthetic navigational structures, where *intentional chance* (our intention, their chance) and simultaneous narratives can create new interpretative cinematic experiences.

CONCEPTUAL DESIGN OF HYPERCAFE
This section offers a detailed discussion of the design, navigation, and linking mechanisms in *HyperCafe*.

Design Aesthetic
HyperCafe has been envisioned primarily as a cinematic experience of hyper-linked video scenes. The video is shown in black and white to produce a film-like grainy quality. In *HyperCafe*, the video sequences play out continuously, and at no point can they be stopped by actions of the user. The user simply navigates through the flow of the video and links presented. This aesthetic constraint simulates the feeling of an actual visit to a cafe where the "real-time video" of the world also plays out continuously. A minimalist interface is employed by utilizing few explicit visual artifacts on the screen. All the navigation and interaction is permitted via mouse movement and selection. For instance, changes in the shape of the cursor depict different link opportunities and the dynamic status of the video. By minimizing the traditional computer-based artifacts in the interface and retaining a filmic metaphor, we hope to provide the user with a greater immersion in the experience of conversations in the cafe. Specific instances of utilizing an intentional design aesthetic in *HyperCafe* will be discussed further.

Navigation and Structure
When the user first enters the *HyperCafe*, a overview shot of the entire scene is revealed, with a view of all the participants and tables in the cafe. The low hum of voices can be faintly heard, suggesting conversation. Moving text at the bottom of the screen provides subtle instructions. The camera moves to reveal each table (3 in all), allowing the user 5-10 seconds to select any conversation (Figure 1). The video of the cafe overview scene plays continuously, forwards and then backwards, until the user selects a table (the audio here is distinct from the video, and remains a constant hum of conversation). Once a choice is made, the user is placed in a narrative sequence determined by the conversations in the selected table. The user may be returned to the main cafe sequence at the end of the conversations (this is but one possibility); a specific conversation may trigger other related narrative events, that the user can choose to follow.

Figure 1: As the camera continually pans across the cafe, many opportunities exist to select a single table of conversation and navigate to the related video narratives.

The user navigates through the logical structure of *HyperCafe* via a hierarchy of video scenes creating linked narrative sequences. At the top level, the main cafe sequence provides access to all other possible narratives. At the second level, conversational narratives for each table are available. A single conversation is chosen randomly if a particular table is selected. The third level consists of a sequential stream of video scenes (representing a conversation) with links to other video conversations in the current or in different tables (i.e., links within or across the hierarchical nodes). Additional aspects of hypervideo structures will be comprehensively discussed in the section "Framework for Hypervideo."

LINK OPPORTUNITIES IN HYPERCAFE
HyperCafe employs several different types of linking mechanisms, both aesthetically and programmatically.

Temporal Link Opportunities
With *HyperCafe*, we offer a departure from hypertext, as it has been presented thus far in popular literature, with our temporal realization of link opportunities. A notable exception of a hypertext with temporal links is *Dreamtime* [25] by Stuart Moulthrop. As we mentioned earlier, previous work [26][9] stresses time attributes and temporal schedules for Hypermedia rather than aesthetic navigational support of temporal opportunities. Traditional hypertext presents users with several text or

image-based links simultaneously, and opportunities in any one node are available concurrently. In narrative situations where we have employed temporal linking, the story proceeds sequentially and opportunities are presented temporally in the form of one or more previews of related video scenes that dynamically fade-in, determined by the context of the events in the current video scene (Figure 2). The user is given a brief window of time (3-5 seconds) to pursue a different narrative path. If the user makes a selection the narrative is redirected to a labyrinth of paths available within the structure of *HyperCafe*, otherwise the predetermined video sequence continues to play out.

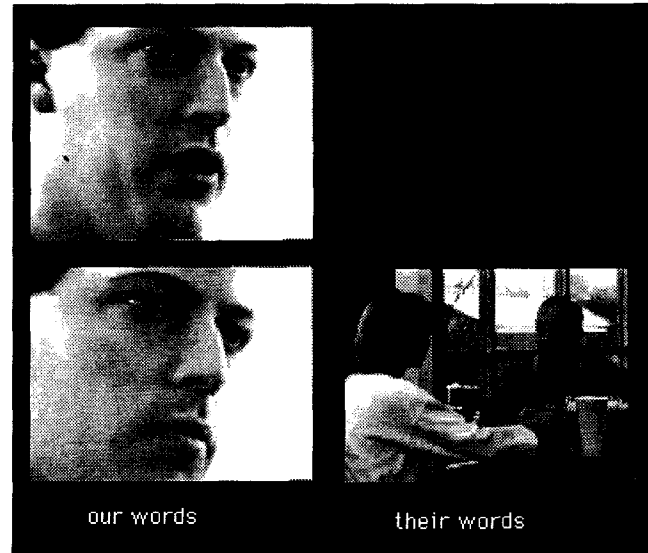our words   their words

Figure 2: As one conversation is shown (video of man on the bottom left), two new temporal opportunities briefly appear (on the top and right) at different points in time. One of the new conversations can now be selected (within a time-frame) to view the related narrative, otherwise they will both disappear.

In some scenes, alternative camera angles can be selected to change filmic perspective and view conversational reactions. Presence of alternative points of view could be shown via camera icons that dynamically appear within the frame of the video, and point in the appropriate direction of view.

Spatial Link Opportunities
In some video scenes, the user can explore the filmic depth of the scene to reveal spatial link opportunities that can trigger other video sequences, such as conversations in a specific table in the background (Figure 3). Such *opportunities* are found *within* the frame itself, where spatial positioning of the conversants in time, recalls or uncovers related interactions, when activated. These spatial opportunities are implemented as dynamically available (transparent) objects

associated with specific aspects in the video frame. With the recent progress in video segmentation and content analysis techniques, it is conceivable that objects in the frame of moving video could be automatically detected and tracked in real-time. For a good discussion of approaches to parsing and abstraction of visual features in video content, see the recent work by Zhang et al. [32].

Generally, we feel that *HyperCafe* requires a more exploratory form of interaction to effectively utilize spatio-temporal links. The user could be made aware of the presence of spatial link opportunities by three potential interface modes: flashing rectangular frames within the video, changes in the cursor, and/or possible playback of an audio-only preview of the destination video when the mouse is moved over the link space. Several large and overlapping rectangular frames could detract from the aesthetic of the film-like hypervideo content, yet the use of cursor changes alone requires the user to continuously navigate around the video space to find links. The cursor-only solution is similar to that utilized by Apple's *QuickTime VR* interface for still images, yet navigation in hypervideo is complicated by moving video (and hence dynamic link spaces). One solution is to employ a combination of modes, where the presence of spatio-temporal links is initially provided by an audio preview in a specific stereo channel to indicate general directional cue. Then changes in the cursor prompt the user to move around the video space, and actual link spaces are shown with a cursor change, coupled, if necessary, with a brief flash of a rectangular frame around the link space. Overlapping link spaces are shown only temporally, and not simultaneously. We are still evaluating which mode or combination of modes is best suited for spatio-temporal links.

Figure 3: The main video narrative (on the left) shows a table with two men in the background. A spatio-temporal opportunity in the filmic depth of the scene triggers another narrative.

Spatio-temporal links also permit the user to select the background of some scenes to return to the main cafe sequence. An "exit space" here allows the user to leave *HyperCafe*.

It must be noted that the destination of temporal or spatio-temporal links may either be entire video scenes or particular frames within a video scene.

Interpretative Textual Links

Textual narration that is annotated to specific video scenes and links between scenes constitutes interpretative text. Such text appears at specific times while the associated video is being played. Longer lines of text appear scrolling horizontally across the bottom, with the directional movement dynamically controlled by the location of the cursor. The text represents associative links based on related discussions of the conversants among different tables. In some cases the text simply represents random bits of the dialog or even the *actual script* of the narrative, sharing the same space as the production video. Text intrudes on the video sequences, to offer commentary, to replace or even displace the videotext. Words, spoken by the participants are subverted and rewritten by words on the screen, giving way to tension between word and image. Traditional hypertext links also permit navigation to related scenes of the videotext.

Figure 4: A video collage or "simultaneity" of multiple colliding narratives, that produce other related narratives when two or more video scenes semantically intersect on the screen.

Several aesthetic effects enhance the reading of the videotext. A video wall of conversations permits users to activate different segments of the videos, dynamically joining them together into *new* conversations. A "collision space" allows users to drag moving video (Figure 4) or scrolling lines of text. Previous work at the Interactive Cinema Group (MIT Media Lab) [12] spatially organized video clips, using the *Collage* notepad, as a way of understanding their content and relationships. We propose a dynamic representation of text and video where the intersection of the "image" or text would trigger other video images and textual narratives. This creates a simultaneity of multiple

"closed" or intersecting narratives [29] along spatio-temporal dimensions. Such interaction between the user and the content produces an entirely new "videotext."

FRAMEWORK FOR HYPERVIDEO

We first summarize definitions of some useful terms before we present a broader framework for hypervideo.

Scene: The smallest unit of hypervideo, it consists of a set of digitized video frames, presented sequentially.

Narrative sequence: A possible path through a set of linked video scenes, dynamically assembled based on user interaction.

Temporal links: A time-based reference between different video scenes, where a specific time in the source video triggers the playback of the destination video scene.

Spatio-temporal links: A reference between different video scenes, where a specified spatial location in the source video triggers a different destination video at a specific point in time.

Temporal link opportunities: Previews of destination video scenes that are played back for a specified duration, at specified points in time during the playback of the source video scene.

Spatial link opportunities: A dynamic spatial location in the source video that can trigger destination videos if selected.

Now we can describe the possible range of structures that can be formed within a general framework of hypervideo. While the Amsterdam Hypermedia Model [14] provides a more generalized approach to the representation and synchronization of media data types, our framework provides a more specific structure for presenting narrative sequences using hypervideo. Overall, it is our intent that a hypervideo structure be able to embody sufficient definition and abstraction to permit the creation and navigation of a network of hyper-linked video scenes. The scenes may involve conversations between multiple participants, distributed within the space or time of the video narratives. It should be possible for the user to explore the narrative spaces of the hypervideo while allowing the developer/author sufficient control of the system to create desired narrative and aesthetic effects.

At the lowest level of the hypervideo, frames of digitized video[1], are assembled into logical units called scenes. An example of a scene may be a character in the narrative walking to a door, opening it and exiting

the room. Scenes themselves are assembled into larger structures which are displayed "end to end" to form narrative sequences as shown in Figure 5. There are no restrictions on either the size of scenes or narrative sequences, or the number of scenes that make up a narrative sequence, except that there must be at least one scene in each. However, scene connections can also embody contextual information present in the narrative, allowing multiple destinations *(multivalent links)* as a scene plays out. Such contextual information permits decisions based on chance (for a random selection of the next scene within the context), the number of previous visits to that scene, or whether or not the user has previously visited other scenes in the "video space." These types of "decision points" in the narrative are closely related to Zellweger's *programmable paths* and *variable paths* [31].

Implicit in this definition of scene connections is that narrative sequences may "share" scenes as demonstrated in Figure 6. Here, scene 7 is utilized in both narratives and the decision whether scene 8 or scene 12 follows is based on the context in which the choice is evaluated. Note that this usage is similar in some respects to Michael Joyce's hypertext fiction, *afternoon, a story* [19]. A single node in *afternoon* can assume multiple meanings if encountered in different contexts. While the text in the node itself appears the same, context has altered the meaning.
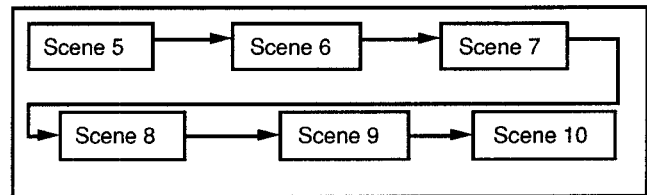

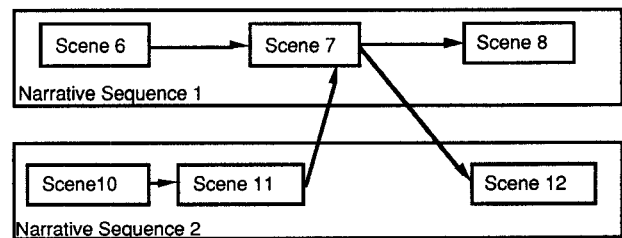**Figure 5: Linear Narrative Sequence of Scenes**

Figure 6: Scenes Shared between Narrative Sequences

Throughout the hypervideo, decisions about the composition of narrative[2] sequences are being made not

---

[1] For convenience of our definition, we will assume that video frames carry with them an audio component recorded concurrently with the video frame.

[2] Narrative, by definition, is traditionally confined to linear media. "Narrative," as we are employing it, refers to one thread of many in a weave of discursive trajectories. With hypervideo, interpretative narratives are dynamically assembled by the user's interaction with the videotext.

just by the developer, but also by the user. As the author defines the sequence of scenes in the narrative and their structured relationship to each other, he designates some of them as link opportunities for the user. (In general, the number of link opportunities will be substantially smaller than the number of links connecting scenes, as most scenes are fairly brief and are simply linked to the next scene of a narrative for continuity.)

One can imagine that a hypervideo could be defined simply as a graph of nodes and links, where the nodes of the video correspond to our scenes. This would be equivalent to removing the outer boxes in Figure 6 which represent the narrative sequences. While this definition might be sufficient, it fails to convey a significant semantic aspect of the structures. We consider the body of narrative sequences a "map," overlaid onto the graph structure of the hypervideo to provide authors with a familiar (and higher-level) concept. The narrative sequences in our formulation of hypervideo is similar to the writing spaces in *Storyspace* [30] or the "overviews" used in Brown University's Intermedia system [21].
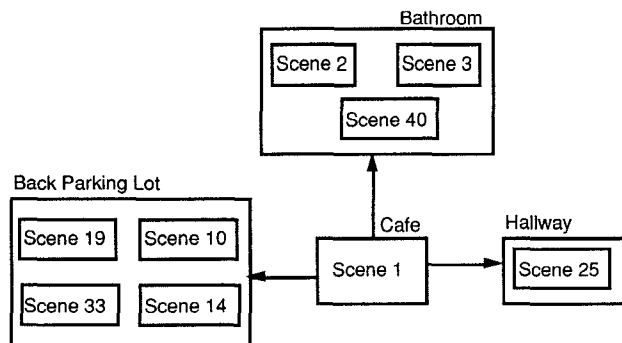

Figure 7: Spatial Map of Narrative Sequences

It may be desirable to define other semantic maps to provide alternative frameworks for the hypervideo structure. An example might be a spatial map to conceptually guide authors in relating space to the structure of scenes (via the spatial link opportunities). The example in Figure 7 shows how each of the exits of a cafe could result in particular video sequences. Navigational bridges, consisting of specific camera pans around the cafe, are utilized between the "narrative video spaces." This abstract spatial map allows the author to more easily visualize the fictive space of his or her creation.

INTERSECTION OF HYPERTEXT AND FILM/VIDEO

In this section, we discuss the influences of film and hypertext to our work, and the implications of this new discursive formation, hypervideo.

HyperCafe and Hypertext

HyperCafe engages Jay David Bolter's notion of "topographic writing," that is, "writing *with* places, spatially realized topics" [7]. While *HyperCafe* is not an electronic writing environment, the spatiality and placement of text and video on the screen are vital to the user's experience. The user creates his or her own "videotext" artifact with *HyperCafe*. The nature of *HyperCafe*'s video interaction is, in Michael Joyce's words, exploratory (as opposed to constructive) [18]— i.e., users cannot add their video work to the body of *HyperCafe*. This constraint is largely imposed by the media we are working in: digital video takes time to produce, from shooting to digitizing, to the disk storage space required. However, in future prototypes we would like to make provisions for the user to his or her own text in selected spaces, and to link it to video clips and textual interactions. While the Cafe isn't yet a realized constructive hypertext, it is desirable that the user be able to save their links through the program, providing for them a memory (or mis-memory) of their encounters, should they wish to return to the program and recover, replay, or rewrite those encounters. They are, in Joyce's words, "versions of what they are becoming, a structure for what does not yet exist" [18].

As the user moves through conversations and makes choices, the spatial and narrative contexts necessarily shift: videos play in different portions of the screen or concurrently, suggesting relationships between the clips based on proximity, movement, and absence; text appears and disappears, ghosted annotations and mock dialogue-revisions. These shifts and events appear based on user interaction, or are intentionally hidden. The same clip may play during a "car crash" narrative line as would play during a "do you remember me?" narrative sequence. The clip stays the same—the context changes. By recontextualizing or repositioning identical clips at several points in the program, we are shifting the meanings of our media, asking the user to engage in building the text and context, making meaning.

In *HyperCafe*, there is an inherent determination to make all chance encounters of the videotext meaningful. The navigation is thus always "contingent" and the reading is subject at every moment to "chance alignments and deviations that exceed the limits of any boundaries that might be called 'context'." [16]. J. Yellowlees Douglas, in charting the "narrative of possibilities" of *afternoon, a story*, describes the experience of visiting the same space four times and not realizing the words were the same, that only the context had changed [11]. Douglas uses *afternoon* and *WOE*, also by Joyce, as examples of Umberto Eco's concept of the open work, or a work whose possibilities even on multiple readings are not exhausted. When the user's session with *HyperCafe* ends, contingencies remain, based on "indeterminabilities operating between the

gaps of the reading" [16], leaving behind the possibility of an unexhausted, if not inexhaustible, text.

HyperCafe and Film

HyperCafe is not entirely textual, or even primarily textual—it communicates most of its information to the user with digital video. We have attempted a marriage of film and hypertext, whose properties (in avante garde film, at least) are already quite similar.

HyperCafe implicates cinematic form as one its models/modes of representation. Our digital video clips are composed entirely of head and body shots of actors, and of movement between them. Close-ups are used to convey a sense of emotion or urgency, and panning long shots—establishing shots—are used to set the speakers up, showing them in relation to one another. However, another process of signification is at work with our choice of long shots. They allow users to navigate between actors and between stories. When actors are in view at a point where a link space exists, the user may choose to move there. Thus, the pan signifies differently in *HyperCafe* than it would on a movie screen. A detached pan becomes an opportunity for action, serving as a *navigational bridge* to and between narratives.

Instead of complying with the typical shot/reverse shot style of representing conversation in film, *HyperCafe* allows for a new grammar to be defined. Reverse shots can be answered by an actor on the other side of the screen, from an entirely different clip. Textual intrusion in the space between could disrupt or enable the conversation. The mise-en-scene of the computer screen, then, can be defined in wholly different terms. Shot composition is no less important in hypervideo, but the choice of elements with which to compose are quite different than in film.

Just as film form structures the user's interaction with the text by the filmmaker's choice of shot, lighting, actor's delivery of lines, and cinematography, so too does the hypervideo interface structure a reader's formal interaction with the videotext. Here, a user's assumptions about computing environments and existing hypermedia applications will structure his or her assumptions about interaction, assumptions which will invariably prove different than that user's assumptions while watching a movie.

While we are clearly departing from traditional film and video form, questions must be asked: does *HyperCafe* reveal power/authority/univocal traits of traditional cinema? What new (if any) forms of representation are we enabling? These questions must necessarily be answered by extended evaluation and analysis by these authors and other interested theorists. Hypervideo is at its infancy; there exists a unique window of opportunity today to invent this discipline, to move away from traditionally confining modes of representation, to support a more open, collaborative body of work.

CONTENT PRODUCTION AND PROTOTYPE DEVELOPMENT

The content for *HyperCafe* consists primarily of digital video, taped with two Super-VHS video cameras (HyperCams 1 and 2). External microphones captured individual conversations and ambient room noise. The two cameras simultaneously provided two different perspectives/shot angles for each scene. One camera generally remained stationary, providing a long shot, while the other was mobile, providing close-ups and movement within the shot. Some extreme close-ups (like actor's lips) provided the desired dramatic effects for particular conversations. Several top-level and *below-level* pans (shooting the actors' feet) were taken to provide *navigational bridges* between tables.

After the video shoot, all the video scenes (over 3 hours) were edited, manually logged and transcribed into *Storyspace* [30]. A linear thread through the *Storyspace* document was created and later other interpretative hyper-links added (Figure 8). *Storyspace* served as a powerful tool for assessment of the video scenes and greatly aided in the editing and selection of appropriate scenes for use in the digital version. The *Storyspace* "hyper-script" was also utilized to create and simulate (in text only) the multiple narrative sequences through the digital video. It must be noted that in *HyperCafe*, we did not utilize the potentially rich semantic space of the spatial hypertext nodes [24]; our screen mapping of the conversations closely followed the general structure of the writing spaces in *Storyspace* and the actual layout of the Cafe. It would be interesting to consider the effect of overloading the arbitrary or intentional semantics of the writing spaces to influence the *architectonic* [20] presentation of the hypervideo.
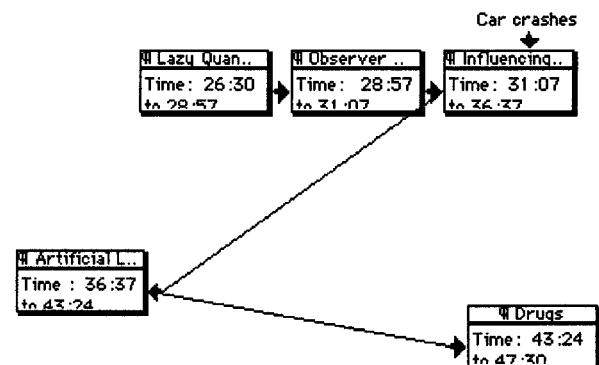

Figure 8: A portion of the HyperCafe script organized and hyperlinked using Storyspace

Selected video scenes were captured on a Macintosh PowerPC 8100/80 with *Adobe Premiere* 4.2 [1] using a 160x120 resolution at 15 frames per second (fps). A black/white filter was applied to the video producing a film-like grainy quality. A total of 25 minutes of video was eventually captured, occupying 300 MB of disk

7

space. These files were segmented into 48 QuickTime movie files, and compressed via a Cinepak codec, reducing the combined size of the files to 150 MB. *Macromedia Director* 4.04 on the Macintosh [23], was selected as the development platform for the initial prototype. Director was utilized due to its ability to control digital video and for rapid prototyping of the design concepts. All the QuickTime movies were imported into Director's multimedia database[3]. Ambient restaurant sounds and interpretative text were added for use in several narrative sequences. Lingo scripts[4] were written to provide interactivity and hyper-linking. The completed Director movie was compiled into a projector (movie player) that could be played back independently on the Mac or PC platform.

TOWARDS A HYPERVIDEO TOOL

Having developed a proof-of-concept prototype in Macromedia Director, we believe that a broader software tool can be developed for hypervideo authoring and navigation. The tool could function similarly to *Storyspace*, in that it would permit placement of (pointers to) digital video and textual content into a hierarchy of hyper-linked nodes. Several navigational paths through these video nodes could be authored. Time attributes should be integrated in the hypervideo model and manipulated at a higher level [9] such that temporal links can be synchronized with playing video. In the navigation mode, the software would dynamically generate link opportunities and permit multiple paths through the video text.

Such a tool should aid both in the pre-production and post-production of the hypervideo product. In pre-production, when no video has been shot, the tool could be used to write hypertextual scripts (in place of the video), and thus build the initial linked hierarchy of the videotext. The hypertextual scripts would later aid in editing and selecting appropriate video scenes. In post-production, after all the video has been shot, edited, and captured, the video scenes would be simply placed in the appropriate nodes. Additional interpretative text, temporal, and spatio-temporal links could be added. It must be noted that temporal links are explicitly added by the authors, based on their creative intentions about the videotext. These links later become "opportunities" that are generated for the user as the videotext is navigated, based on the state of the user in the system (see MIT's *ConTour* system [10], for a more formal approach to dynamic generation of media elements). The videotext could be navigated or "read" at any time in the production. In fact, the videotext would never be

---

[3]It must be noted that only the pointers, with references to the digital video resources, are stored in the database, and not the actual video. This minimizes the size of the compiled projector, thus permitting several versions to be independently distributed based on existing video resources, perhaps on CD-ROM.

[4]Lingo is Director's proprietary object-oriented programming language.

complete since each "reader" may be allowed to add their own interpretative text, voice annotations, and links to continually expand the hypertext and overall link space. Users could create several personalized interpretations of the videotext or collaborate to develop a single multi-user videotext.

HyperCafe demonstrates an application of hypervideo for the production of fictional narratives. This could be utilized by writers, film and video producers to develop non-sequential, hyper-linked video narratives. Layers of textual interpretations can be added to these narratives enabling richer experiences. Different characters could be utilized to drive the narrative along varying paths, such as in *Agent Stories* [3]. We hope that this framework sets the stage for new forms of creative expression, where the interaction and navigation through the videotext itself creates the aesthetic and personalized interpretations.

A hypervideo tool could also be utilized by filmmakers producing conventional films. It would offer the filmmaker alternate structures of conceiving and creating a film than would otherwise be possible without such a tool. For instance, dynamically previewing alternate takes and sequences are possible with hypervideo; sequences can be mapped out and rearranged, much as nonlinear digital video editing suites enable video editors to do. However, with a built-in hypertext tool in the package, a hypertext script could be linked to the video sequences, and a hypervideo tool could be used to makes hotspots in the video frame, or link related text to a particular point in a video sequence. The entire process of conceiving a film script to editing a film, even a linear film, could potentially be reorganized and integrated using a hypervideo tool.

FUTURE WORK

For any hypermedia tool, issues regarding scalability eventually become important and sometimes problematic. A reasonable number of links from one video node to another should be supported. If at a point in time more temporal links exist than can fit on the screen, thumbnails should be automatically generated or the number of simultaneous temporal links constrained. Another concern is the possibility of "dead ends" to the continuously playing video narrative (what Terry Harpold considers the "moment of the non-narrative" [17]). One of the aesthetic goals in *HyperCafe* was to never permit a moment where the video would stop and break the cinematic experience of the user. Yet as the nodal structure of the videotext grows more complex, the authors must painstakingly ensure that all video sequences lead to other sequences, and thus all nodes are *non-terminating*. A suitable approach needs to be developed to ensure that any terminating nodes are returned to either a prior node, the next ordered node after the prior node, or a parent node. In addition, specific "filler" sequences could be shot and played in loops to fill the "dead-ends" and holes in the narratives.

During the video production, innovative camera techniques aided in producing "navigational bridges" between some scenes without breaking the filmic aesthetic. Movement between tables in the cafe was enabled by triggering specifically shot video segments between the tables. We can point to one recent attempt to produce continuous digital video navigation in the *CityQuilt* [13] program by Tritza Even. In *CityQuilt*, Even employed digital video techniques in *Adobe Premiere* to provide panning within moving video, permitting a user to navigate across an endless canvas of video scenes of New York city. We believe that further development of such analog and digital video production techniques is essential to effectively address the navigational issues in hypervideo.

Collaborative hypervideo would require distributed video available over networks with suitable bandwidth. In 1993, the film *Wax or the discovery of television among the bees*. was multicast via the MBone (multimedia backbone) of the Internet to about 450 sites. Later "Waxweb" [6] was developed as an experimental hypertext groupware project. This indicates that a system permitting hypervideo (authoring and navigation) over the Internet would be desirable. If all the video content is made available on CD-ROM, then it is feasible that multiple versions of the "videotext structure" (i.e., player files) could be accessed and modified by users over the Internet (via WWW or MOO), permitting collaborative authoring of hypervideo. In a multi-user environment, appropriate locking schemes would ensure data integrity if users are permitted to add their own links or text annotations. Some form of version control may allow users to access all previous evolving links and annotations, and hence "navigate temporally" through the *historic state* of the videotext.

The current system provides navigation of hypervideo in a two-dimensional space. Previously, tools have been developed to present motion picture time as a three-dimensional block of images flowing away in distance and time [12]. For hypervideo, a three-dimensional space would permit representation of multiple (and simultaneous) video narratives, where the added dimension could signify temporal or associative properties. Video sequences that have already been visited could be represented by layers of (faded) video receding into the background. Additionally, a series of forthcoming links and prior (unrealized) links could also be shown in the foreground/background of the three-dimensional space. A 3D authoring and navigation tool for hypervideo, perhaps called "VideoSpace" (not unlike *Storyspace* for hypertext), could be envisioned for several interesting applications, such as multi-dimensional hypervideo information spaces.

CONCLUSION
The hypervideo system described in this paper provides a glimpse into the potential of creating dynamic hyper-linked videotext narratives. An aesthetic design of navigation and structural representation permits a new form of videotext expression for authors, and interpretative experiences for readers. *HyperCafe* is unique in that it presents "aesthetic opportunities" for navigating to linked narratives, primarily on a spatial and temporal basis. The presence of temporal and spatio-temporal link opportunities creates a new grammar for hypermedia applications, based on a cinematic language. In *HyperCafe* the textual narratives intersect with dynamic video sequences, producing interpretative videotexts. The techniques and methodologies for navigation in the "video space" are experimental and untested, yet provide an appealing minimalist approach. We hope that other developers and theorists will consider the definition and navigational structures we have mapped for hypervideo in constructing their own hypermedia applications and creative expressions.

REFERENCES
1. Adobe Premiere 4.2, a non-linear digital video editing and production software, distributed through Adobe Systems, Inc. 1995.

2. Arons, Barry. "Hyperspeech: Navigating in Speech-Only Hypermedia," *Proceedings of Hypertext '91*, ACM, December 1991, pp. 133-146.

3. Beacham, Frank. "Movies of the Future: Storytelling with Computers," American Cinematographer, April 1995, pp. 36-48.

4. Bernstein, Mark, Jay David Bolter, Michael Joyce and Elli Mylonas. "Architectures for Volatile Hypertext," *Proceedings of Hypertext '91*, ACM, 1991, pp. 243-260.

5. Bers, Joshua, Sara Elo, Sherry Lassiter and David Tamés. "CyberBELT: Multi-Modal Interaction with a Multi-Threaded Documentary," *Proceedings of CHI '95*, ACM, 1995, pp. 322-323.

6. Blair, David. *Waxweb*, a hypertext groupware project at Brown university, based on David Blair's film *Wax*. URL: http://bug.village.virginia.edu/

7. Bolter, Jay David. *Writing Space: The Computer, Hypertext, and the History of Writing.* Lawrence Erlbaum and Associates, Hillsdale NJ. 1991.

8. Brøndmo, H.P. and Glorianna Davenport. "Creating and viewing the *Elastic Charles* - a hypermedia journal," in McAleese R. and Green, C. (eds.) *Hypertext: State of the Art,* Oxford: Intellect, 1991, pp. 43-51.

9. Buchanan, M. Cecelia and P.T. Zellweger. "Specifying Temporal Behavior in Hypermedia Documents," *Proceedings of Hypertext '92,* ACM, 1992, pp. 262-271.

10. Davenport, Glorianna and Michael Murtaugh. "ConText: Towards the Evolving Documentary," *Proceedings of Multimedia '95,* ACM, 1995, pp. 381-389.

11. Douglas, J. Yellowlees. "Understanding the Act of Reading: the *WOE* Beginners' Guide to Dissection.," *Writing on the Edge,* 2.2. University of California at Davis, Spring 1991, pp. 112-125.

12. Elliot, Eddie and Glorianna Davenport. "Video Streamer," *Proceedings of CHI '94,* ACM, 1994, pp. 65-66.

13. Even, Tritza. "CityQuilt: A Navigable Movie," *Proceedings of Multimedia '95,* ACM, 1995, pp. 279-280.

14. Hardman, L., D.C.A. Bulterman, and G.V. Rossum. "The Amsterdam Hypermedia Model: Adding time and context to the Dexter model," *Communications of the ACM,* 37(2), February 1995, pp. 50-62.

15. Hardman, L., D.C.A. Bulterman and G.V. Rossum. "Links in Hypermedia: the Requirement for Context," *Proceedings of Hypertext '93,* ACM, 1993, pp. 183-191.

16. Harpold, Terence. "The Contingencies of the Hypertext Link," *Writing on the Edge,* 2.2. University of California at Davis, Spring 1991, pp. 126-138.

17. Harpold, Terence. Personal conversions about HyperCafe.

18. Joyce, Michael. "Siren Shapes: Exploratory and Constructive Hypertexts," *Academic Computing 3* (4), 1988, pp. 10-14, 37-42.

19. Joyce, Michael. *afternoon, a story.* Computer disk. Cambridge, MA: The Eastgate Press. 1987.

20. Kaplan, Nancy and Stuart Moulthrop. "Where No Mind Has Gone Before: Ontological Design for Virtual Spaces," *Proceedings of Hypertext '94,* ACM, 1994, pp. 206-216.

21. Landow, George P. *Hypertext: The Convergence of Contemporary Critical Theory and Technology.* The John Hopkins University Press, 1992.

22. Liestøl, Gunnar. "Aesthetic and Rhetorical Aspects of Linking Video in Hypermedia," *Proceedings of Hypertext '94,* ACM, 1994, pp. 217-223.

23. Macromedia Director 4.04, a cross-platform authoring tool for multimedia authoring, distributed through Macromedia, Inc. 1995.

24. Marshall, C. Catherine and Frank M. Shipman III. "Spatial Hypertext: Designing for Change," *Communications of the ACM,* 38(8), August 1995, pp. 88-97.

25. Moulthrop, Stuart. *Dreamtime,* A hypertext with time-constrained hyperlinks, 1992. URL: http://raven.ubalt.edu/Moulthrop/hypertexts/dreamtime.html

26. Ogawa, Ryuichi, Eiichiro Tanaka, Daigo Taguchi and Komei Harada. "Design Strategies for Scenario-based Hypermedia: Description of its structure, Dynamics, and Style," *Proceedings of Hypertext '92,* ACM, 1992, pp. 71-80.

27. Ogawa, Ryuichi, H. Harada and A. Kaneko. "Scenario based Hypermedia: A Model and a System," in A. Rizk, N. Streitz and J. Andre (eds.), *Hypertext: Concepts, Systems, and Applications,* Cambridge University Press, 1990, pp. 38-51.

28. Potts, C., Jay David Bolter and A. Badre. (1993) "Collaborative Pre-Writing With a Video-Based Group Working Memory," *Graphics, Visualization, and Usability Center* (Georgia Institute of Technology, USA) *Technical Report 93-35.*

29. Rosenberg, Jim. "Navigating Nowhere/Hypertext Infrawhere," ECHT 94, *ACM SIGLINK Newsletter,* December 1994, pp. 16-19.

30. Storyspace 1.3w, a hypertext writing environment developed by Jay David Bolter, Michael Joyce, and John B. Smith. It is distributed through Eastgate Systems Inc., 1995.

31. Zellweger, P.T. "Scripted Documents: A Hypermedia Path Mechanism," *Proceedings of Hypertext '89,* ACM, 1989, pp. 1-14.

32. Zhang, H.J., C.Y. Low, S.W. Smoliar and J.H. Wu. "Video Parsing, Retrieval and Browsing: An Integrated and Content-Based Solution.," *Proceedings of Multimedia '95,* ACM, 1995, pp. 15-24.