

Combining Spatial and Navigational Structure in the Hyper-Hitchcock Hypervideo Editor

Frank Shipman, Andreas Girgensohn, Lynn Wilcox
FX Palo Alto Laboratory, Inc.
3400 Hillview Avenue
Palo Alto, CA 94304, USA
+1-650-813-{7244, 7574}

{shipman, andreasg, wilcox}@fxpal.com

ABSTRACT

Existing hypertext systems have emphasized either the navigational or spatial expression of relationships between objects. We are exploring the combination of these modes of expression in Hyper-Hitchcock, a hypervideo editor. Hyper-Hitchcock supports a form of hypervideo called “detail-on-demand video” due to its applicability to situations where viewers need to take a link to view more details on the content currently being presented. Authors of detail-on-demand video select, group, and spatially arrange video clips into linear sequences in a two-dimensional workspace. Hyper-Hitchcock uses a simple spatial parser to determine the temporal order of selected video clips. Authors add navigational links between the elements in those sequences. This combination of navigational and spatial hypertext modes of expression separates the clip sequence from the navigational structure of the hypervideo. Such a combination can be useful in cases where multiple forms of inter-object relationships must be expressed on the same content.

Categories and Subject Descriptors

H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems – *video*. H.5.4: Hypertext/Hypermedia – *navigation, user issues*.

General Terms

Design, Experimentation, Human Factors.

Keywords

Spatial hypertext, Hypervideo, Interactive Video.

1. INTRODUCTION

Spatial hypertext emerged as an alternative to the explicit expression required when authoring links in navigational hypertext. In a spatial hypertext, most expression occurs via the application of visual attributes (e.g. color, border width) or spatial layout (e.g. lists, stacks). We are exploring the use of both implicit spatial and explicit relational modes of expression within Hyper-Hitchcock, a hypervideo editor. Hyper-Hitchcock uses a simple spatial parser to determine the sequencing of video clips placed in a two-dimensional workspace. Navigational links from video elements to additional video are explicitly attached to elements in the workspace.

Hypervideo research has tended to focus on the characteristics of hypervideos [2] or on automatic detection of objects for anchor

tracking [1, 5]. Hypervideo authors have been left in the world of scripting languages with the exception of MediaLoom [6], which allows users to create links between existing video files but does not support video composition. The need to support the composition of segments of linear video and links between these compositions led to a spatial / navigational hybrid approach to authoring.

Integration of spatial and navigational hypermedia mechanisms is not unique to Hyper-Hitchcock. Spatial hypertexts have included links to take the viewer from one location to another location or another spatial hypertext or external document. What is unique is the role of the spatial parser in structuring the resulting hypervideo.

2. HYPER-HITCHCOCK

Hyper-Hitchcock supports a limited form of hypervideo we call “detail-on-demand video” due to its applicability to situations where a person would take a link to view more details on the content currently being presented. Video takes (on/off boundaries in source video) are automatically divided into video clips based on camera motion and lighting. These clips (top-left of Figure 1) are selected by dragging them into the workspace (bottom, Figure 1).

Clips can be grouped into composites, which can be part of even larger composites. Authors can switch to a split-workspace view to

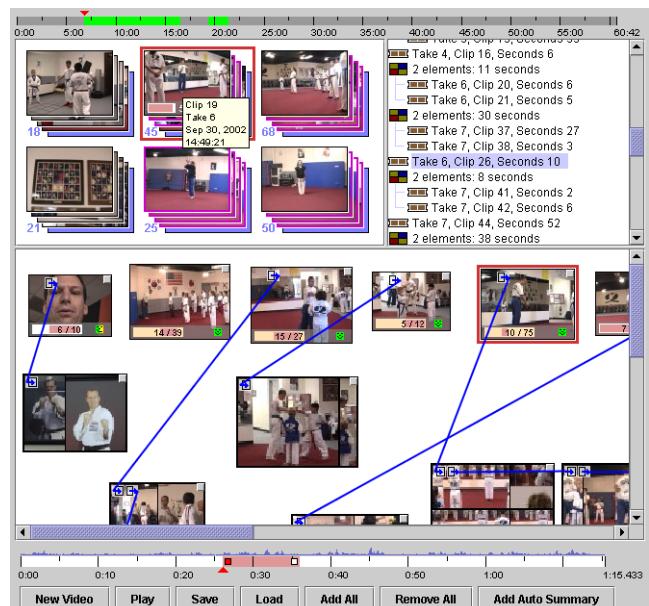


Figure 1: Spatial/navigational structure in Hyper-Hitchcock.

Copyright is held by the author/owner(s).

HT '03, August 26-30, 2003, Nottingham, United Kingdom.
ACM 1-58113-704-4/03/0008.



Figure 2: Editing a hypervideo summary. The left workspace is the automatically generated four-level summary. The right shows the components of one of the composites, allowing the components and their links to be modified.

modify the components of composites, as seen in Figure 2. By rearranging the elements in the workspace on the right, their sequence in the composite is changed. The spatial parser recognizes horizontal/vertical lists and the playback order of the video elements is defined accordingly.

Every level of the hierarchy of video clips can have a single navigational link to another element (clip or composite) in the workspace. When more than one link would be active at a time, the most detailed link (that applied to the lowest level element in the hierarchy) is the one presented during video playback. Links in detail-on-demand video include a destination video sequence, an offset into that sequence, a label for the link, and return behaviors for when the destination completes playback or the user aborts its playback. (See right side of Figure 2.) Return behaviors include returning to the point in the source video where the link was taken, returning to the beginning of the source anchor, returning to the end of the source anchor, and stopping video playback all together. For more information on the detail-on-demand video representation and the Hyper-Hitchcock interface see [4].

Hypervideo can aid the location of specific video segments in long videos. But the human authoring of hypervideo indices is costly. In addition to being an authoring tool for hypervideo, Hyper-Hitchcock includes the ability to automatically generate multi-level hypervideo summaries that allow a viewer to navigate from an overview of the video, to longer summaries of the video, to the full video. Figure 2 shows a four-level automatically generated hypervideo: 30 second, 3 minute, and 12 minute summaries in the top three levels and the full hour of video in the bottom row. Algorithms for generating hypervideo summaries are detailed in [3].

3. DISCUSSION

Hyper-Hitchcock combines navigational links and spatial structures for author expression. By separating the expression of sequencing and linking, it provides an example for other editors where multiple overlapping structures are authored. Hitchcock's automatically generated hypervideo summaries provide an example of the potential to create initial spatial layouts and navigational links that act as a starting point for hypervideo authors.

The spatial parser in Hyper-Hitchcock is limited to list recognition rather than attempting to recognize complicated structures. This is

due to the domain — lists are natural representations of sequences. Recognizing higher-level structures, such as lists of lists could aid the author in terms of scale (making it easier to use the full workspace when composing a sequence). But when the recognized structure and the author's intent do not match, the resulting video would be incorrect and go unnoticed without watching the video.

Detail-on-demand video is a limited form of hypervideo where only one link is available at a time. This allows the potential for playback on devices like DVD players using a remote. Links in detail-on-demand video allow for a variety of author intents. By having a link return to the beginning of its source anchor, the author can provide prerequisite information to the viewer. By having the link return to the end of source anchor, the author provides an alternative perspective or path through the hypervideo. Sequences of links that return to the end of the source anchor can create video indices into content or complex branching structures.

4. REFERENCES

- [1] Hirata, K., Hara, Y., Takano, H., and Kawasaki, S. Content-oriented Integration in Hypermedia Systems, *Hypertext '96 Proceedings*, ACM, New York. pp. 11-21, 1996.
- [2] Sawhney, N., Balcom, D., Smith, I. Hypercafe: Narrative and Aesthetic Properties of Hypervideo, *Hypertext '96 Proceedings*, ACM, New York. pp. 1-10, 1996.
- [3] Shipman, F., Girgensohn, A., Wilcox, L. Creating Navigable Multi-Level Video Summaries. *IEEE International Conference on Multimedia Computing and Expo*, vol. II, pp. 753-756, 2003.
- [4] Shipman, F., Girgensohn, A., Wilcox, L. Hyper-Hitchcock: Towards the Easy Authoring of Interactive Video. *Human-Computer Interaction INTERACT '03*, IOS Press, 2003.
- [5] Smith, J.M., Stotts, D., and Kum, S.-U. An Orthogonal Taxonomy of Hyperlink Anchor Generation in Video Streams Using OvalTine, *ACM Hypertext 2000 Proceedings*, ACM, New York. pp. 11-18, 2000.
- [6] Tolva, J. MediaLoom: An Interactive Authoring Tool for Hypervideo, <http://www.mindspring.com/~jntolva/medialoom/paper> (Accessed May 14, 2003).