

Interfaces for Managing Access to a Video Archive

Andrew Gordon, Smadar Kedar, and Eric Domeshek
Institute for the Learning Sciences
Northwestern University
1890 Maple Ave., Evanston, IL 60201
Tel: +1-847-491-3500
Email: {gordon, kedar, domeshek}@ils.nwu.edu

ABSTRACT

We describe Deja Vu, a video retrieval system which capitalizes on our understanding of the content of the video to provide an effective user interface.

Keywords

Information access, interface design, browsing, search, indexing, retrieval, video archive, visualization.

INTRODUCTION AND MOTIVATION

The dream of the Information Superhighway is now being followed by the nightmare of information access: How do I know what is out there? How can I access what I want when I need it? Previous CHI approaches have focused on managing the complexity of information access by providing meta-structures for visualizing large amounts of information at once and traversing large spaces quickly. However, most meta-structures only take advantage of the form of the data, i.e. its media type or structure, ignoring useful information that can be gleaned from its content. We present Deja Vu, a system designed to enable video producers to retrieve video clips from an on-line stock footage archive. We describe a novel approach for using meta-structures to organize information that capitalizes on our understanding of the *content* of the information. In particular, for stock video clips whose contents are everyday activities, we describe an interface structured around activities and their components.

The novel contribution of Deja Vu to CHI is to extend the application of meta-structures for managing the complexity of information access to structuring interfaces according to the underlying information *content*.

BACKGROUND AND RELATED WORK

Much current effort in accessing large stores of information focuses on meta-structures such as Multiscale viewing [4], viewing information objects at many different scales. The scale can vary by distortion (e.g. fisheye), zooming (e.g. multitrees) or 3D animation (e.g. Butterfly). However, most efforts focus on text or hypertext [5] rather than visual images (with few exceptions such as FilmFinder [1]). In addition, much of the work exploits the form of the data,

not its content. Work in multimedia and AI do address access of video, focusing on particular techniques for video indexing and retrieval [3]). However, on the whole they do not address the CHI issues of visualizing and navigating easily in a large store of video.

ACTIVITIES AS CONCEPTUAL INDICES

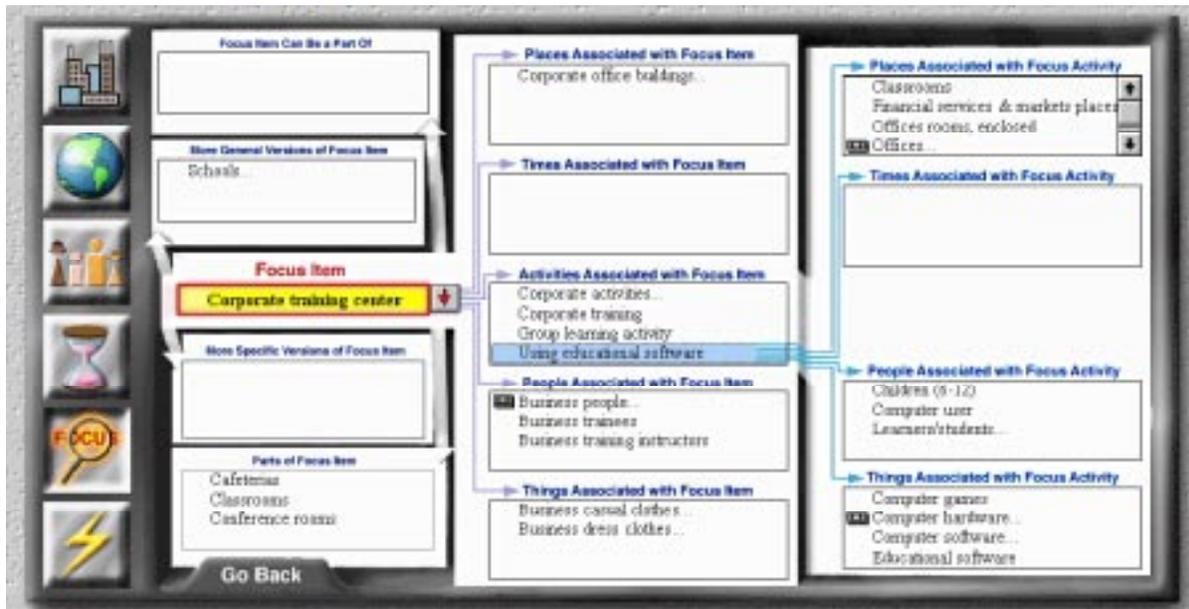
Deja Vu is a system that enables video producers to retrieve video clips from an on-line stock footage archive. Stock video typically consists of scenes of everyday activities which can be reused by video producers to reduce production costs. Given the broad reusability of stock video, it is best organized (indexed) in a way that captures the visualizable content of each clip. Each video clip is indexed using terms to describe its salient features (*conceptual indexing*). We use an extensible vocabulary of everyday activities and the associated people, things, places and times. The organization of this vocabulary has been influenced by artificial intelligence and cognitive theories of memory organization (*Scripts* and *MOPs* [6]) which capture commonsense expectations about the relationship between activities and people, things, places and times experienced in everyday life. By aligning our vocabulary organization with people's natural expectations, we provide a network of terms that can be quickly and easily navigated by the user to browse and search for video.

Although the initial costs of developing a broad indexing vocabulary and manually indexing clips is high, the resulting indices reflect the content more accurately than those which can be obtained through current video analysis techniques, and are less ambiguous than those produced by text-based logging [7].

INTERFACES THAT MANAGE COMPLEXITY

Deja Vu is designed to allow users to search for video clips by browsing through the vocabulary space, using a point-and-click interface. Accordingly, the interface must provide a clear means of navigating the complex vocabulary space in order to formulate the search query, and quickly and accurately locate the needed video clips. We have incorporated a number of interface design principles into Deja Vu that allow us to meet these requirements. We describe them below, illustrating them with an example of how a producer would search for stock footage to be used in a video production on computer-based training.

Zooming: This provides an initial entry point into the large space of terms used for retrieval of video [2]. The zoomer interface provides a global view of all people, all places, all



time, and so on. Within a few clicks the user will have traversed a large space (e.g. all places, cities) and zoomed in on the specific place (e.g. corporate training centers).

Synergy of browsing and search: After the initial zoom, the user browses the space of query terms and selects certain ones to compose as a query (query by navigation [5]).

Browsing organized by activities The browsing interface is organized by activities and their components, reflecting the content of the underlying corpus. A concept zoomed into, such as corporate training centers, is part of an activity, corporate training, which also consists of people (trainees), things (computers), places (classrooms) and time (business day). This organization of the interface is a natural one, and the user should find the concepts they are looking for in a predictable place. As the concepts are found, they are collected to form a query. The browsing interface also enables the user to search for video *related* to what he or she wants by perusing clusters of activities, providing graceful degradation.

Immediate feedback on query: The user may only compose a query of terms which have video clips associated with them. As the query is being composed, the user immediately sees how that affects the retrievable corpus, providing tight coupling [1]. After noticing that there are some clips satisfying the query, the user can view the retrieved set and select the desired clips.

EVALUATION, FUTURE WORK, & CONCLUSION

Deja Vu is implemented in Delphi, running on a Windows PC, and uses a Paradox and Interbase database to manage the video archive and indexing scheme. It currently has 1000 indexed video clips. As part of our iterative design, we performed a formative evaluation on a previous version of the interface with users who have done video production. The users found the system innovative and with potential to

simplify the job of retrieving stock footage, and provided suggestions on the graphic design and labels used in the interface. Our future work includes improvements in the interface design, and field testing. Deja Vu is a proof-of-concept that capitalizes on the content of a large corpus of information to provide interfaces that facilitate information access.

ACKNOWLEDGMENTS

We thank Lannert, Persiko, Reese, Rosenberg, Schmidt, Swanson, and past Deja Vu team members. We also thank Andersen Telemedia for whom this system is designed.

REFERENCES

1. Ahlberg, C. and Shneiderman, B. Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. In Proc. CHI '94 (Boston, April, 1994), ACM press, pp. 313-317.
2. Bareiss, R., and Osgood, R. Applying AI Models to the Design of Exploratory Hypermedia Systems. In the Proceedings of the Fifth ACM Conference on Hypertext, (Seattle, 1993), ACM Press, pp. 94-105.
3. Baudin, C., Davis, M., Kedar, S. and Russell, D. M. Proceedings of the AAAI-94 Workshop on Indexing and Reuse in Multimedia Systems (Seattle, August, 1994), American Association for Artificial Intelligence.
4. Furnas, G. W. and Bederson, B. B. Space-Scale Diagrams: Understanding Multiscale Interfaces. In Proc. CHI '95 (Denver, May, 1995), ACM press, pp. 234-241.
5. Mackinlay, J.D. and Zellweger, P. T. Panel: Browsing vs. Search: Can We Find a Synergy?, In Proc. CHI '95 (Denver, May, 1995), ACM press, pp. 179-180.
6. Schank, R. Dynamic Memory: A Theory of Reminding and Learning in Computers and People. Cambridge University Press, Cambridge, England, 1982.
7. Webber, K. and Poon, A. Marquee: A Tool for Real-Time Video Logging. In Proc. CHI '94 (Boston, April, 1994), ACM press, pp. 58-64.