User Strategies for Handling Information Tasks in Webcasts

Christine Dufour[1]
Université de Montréal
Montréal, QC H3C 3J7
christine.dufour@umontreal.ca

Elaine G. Toms **Jonathan Lewis**
Dalhousie University
Halifax, NS B4H 4H8
{elaine.toms, jmlewis2}@dal.ca

Ron Baecker
KMDI[2], University of Toronto
Toronto, ON M5S 2E4
rmb@kmdi.toronto.edu

ABSTRACT
Webcast systems support real-time webcasting, and may also support access to the stored webcasts. Yet, research rarely examines issues concerning the interface to webcast systems, another form of multimedia system. This paper focuses specifically on how stored webcasts are re-used. Sixteen participants performed three typical information tasks using ePresence, a webcasting system that handles both live and stored video, and contains several tools: a video window, a timeline of the webcast, slides used by the presenter, and a moderator-generated table of contents, that facilitate user access to the intellectual content of a stored video. Use takes place at the level of the webcast, and our analysis assessed user interactivity. The results showed that different types of tasks need different strategies and tools.

Categories & Subject Descriptors: H.5.1 Multimedia Information Systems: Video.

General Terms: Design, Human Factors, Measurement.

Keywords: Search strategies, information tasks, webcasting, ePresence.

INTRODUCTION
A webcast is the Internet audio and/or video stream produced from a live event, or an online simulcast of a broadcast signal [11]. While the act of webcasting was once intended as a live, real-time-viewing-only Internet streaming application, today, many applications exist for which stored webcasts are de rigueur, e.g., online courses, expert presentations, conferences, and training events. The systems used to stream an event are often the same ones used to review stored webcasts, and we questioned this implementation. In this research we examined how a set of typical users interact with stored webcasts, mimicking the process that people use in returning to the record of an event. This study was exploratory, but comprehensive; in this paper we focus specifically on the strategies that users employ for handling information tasks using the webcasting system. By understanding user strategies, we will also understand the effectiveness of webcast interfaces for performing ordinary information tasks.

CHI 2005, April 2–7, 2005, Portland, Oregon, USA.
ACM 1-59593-002-7/05/0004.

PREVIOUS WORK
A webcasting system can be classed as a form of multimedia system, and webcast is a form of video. For the last decade, digital video research has focused on technological issues such as video indexing, capture, compression and storage [9]. More recently, this focus has shifted toward the interface [9] to concentrate on the evaluation of surrogates in task-oriented contexts with a particular emphasis on textual or visual surrogates [2,3,5]. Audio surrogates have received less attention [12]. This current research as reviewed by [6] emerged initially to facilitate access to and use of video repositories such as the Fischlàr Digital Library and the Informedia project.

Very little has been done on webcasting systems. The existing webcasting systems are mostly commercial systems with no publicly available research. Examples of academic research projects include ePresence [1], the Berkeley Internet Broadcasting System (BIBS) [8], and the "Authoring on the Fly" system [4]. These looked at specific tools/surrogates that support the use of stored or live webcasts. As different surrogates are likely needed for different types of video (e.g., informational audio-centric, narrative-entertainment) [7], additional research is needed.

In this research, we conducted a user study to examine use of stored webcasts using three types of typical information tasks. This type of task isolates information chunks, and interprets or manipulates those chunks. In this research, a holistic data collection was employed, one that assessed procedural processes, as well as cognitive processes involved in executing the tasks.

METHODS

System: ePresence
The stored webcast interface of ePresence (http://epresence.tv) was used for this study. In addition to the video window that displays the presenter, the system has several tools to facilitate interaction with the content of webcasts: (1) the slides used in the presentation, (2) a table of contents (toc) created in real time that identifies the 'chapters' in the presentation, (3) a timeline of the presentation (Figure 1). Notably the timeline and toc are synchronized while the slides are under user control.

[1] Research done while a post-doctoral fellow at Dalhousie University and the University of Toronto.

[2] Knowledge Media Design Institute.

Figure 1. ePresence interface for stored webcasts.

Tasks

The information tasks used were derived from [12]:

A. Selection: select from a set of three webcasts those that are relevant to a specific term paper topic;
B. Specific questions: answer three factual questions using a single webcast;
C. Gist: summarize the main theme(s) or gist of a webcast for another person.

The participants had a limited time to fulfill those tasks: 15, 10 and 15 minutes, respectively for tasks A, B and C. The webcasts (approximately 40-45 minutes in length) came from past lectures hosted by the Knowledge Media Design Institute (http://epresence.tv/website_archived.aspx). The webcasts were selected for their interest to the targeted audience, their delivery by well-known scholars, and their potential applicability for one of the tasks.

Participants

The participants were recruited from programs that dealt with the topics covered by the videos used in the study. The sixteen participants (M=10) were young (75% less than 35 years old), and mainly students (88%). About 1/3 had undergraduate degrees and 63% had master degrees. They had used digital video but were not heavy users. Almost half of them had used the ePresence system, and were familiar with the topics of the webcasts, but very few had already viewed those particular webcasts.

Procedure

The study was conducted using a standard desktop computer with 17 inch monitor. Two applications were simultaneously loaded: the ePresence system, and a custom-designed research system adapted from WiIRE [10] for data collection. The latter controlled the study using a series of HTML forms and, directions to handle the flow of the study and forms for handling participants' responses to questions. A participant maneuvered between the two applications using the icons on the task bar. Each session took about 2 hours and followed a protocol similar to [10]:

- After consenting to participate, participants completed a demographic and digital video experience questionnaire;
- Participants had practice time with the system;
- For each of the three tasks, which were assigned in random order, participants:
  - answered questions about their knowledge of the webcast(s) used, and of the webcast topic;
  - worked on the assigned task using ePresence;
  - on task completion, answered questions regarding their perception of the process and the level of accomplishment attained in doing the task;
  - verbally annotated the session using a 'talk after' style while the screen capture of that session was replaying;
- When the three tasks were completed, participants were interviewed for more general information concerning their use and assessment of the ePresence system.

The data were collected using three mechanisms: (1) a Web-based database that captured questionnaire responses, (2) digital audio files for interviews, and (3) video screen-capture software to record the participant's interaction with ePresence while doing the tasks.

Data analysis

To isolate user strategies, the video screen-capture files were manually coded using video editing software, HyperResearch. We created a predefined set of codes that represented all actions taken by the participants, and inserted these codes into the video screen-capture files. This process resulted in a log for each user-task session and contained the action, the order in which each action was taken, and amount of time spent on that action (before another action was taken). In addition, the 'talk-after' interviews were transcribed, and coded for strategies, using Qualrus, a content analysis software tool.

RESULTS & DISCUSSION

Overview

Three types of data were examined to assess search strategies. We examined a) how the interface tools were used by task type, b) how combination of tools were used, c) how participants decided how to use those tools.

Tools usage

All tools were present in some form on the interface (see Figure 1). The search tool for example was represented by a button. The other tools: slide, toc, video window, timeline were always visible. In addition browser buttons for standard web navigation could be employed. Figure 2 illustrates the use of each tool by task measured by the number of times each tool was used. As illustrated, there were significant differences by type of task ($F(2,47)=7.422$, $p<.0001$), as well as significant effects within task by tool use: toc ($F(2,47)=3.277$, $p=.047$), video ($F(2,47)=25.360$, $p<.0001$), search ($F(2,47)=7.268$, $p=.002$), browser actions ($F(2,47)=23.704$, $p<.0001$).

The toc was the tool used most frequently in devising a gist, but was less useful for factual questions; the search button was the exact opposite. Note that the timeline, created specifically for using stored webcasts was the least used tool. The video window, on the other hand, was used for Task A, but much less used for the other two tasks. From

this data, we concluded that different tasks require different tools, even when the tasks seem tightly related.

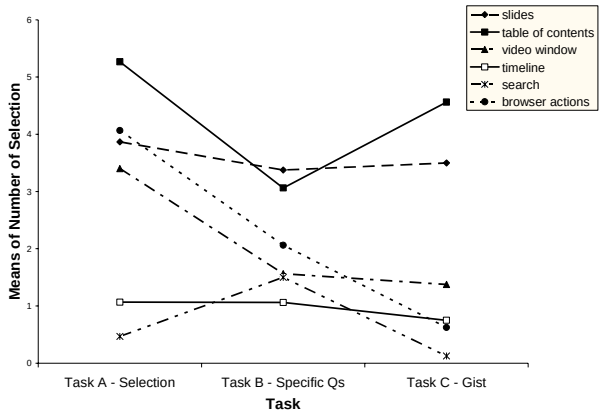

Figure 2. Number of times tool selected by task type.

Combinations of Tool Usage

To further explore the use by tool, we examined patterns of usage using an N-gram pattern technique. Was there a common pattern of tool use among the participants, since there clearly were differences by task? The coded logs were analyzed to extract user strategies:

- The actions were aggregated by interface tools and other types of actions resulting in 5 categories: slides, video, toc, timeline, search, and browser (e.g. using the back button), as well as task-related actions (e.g. going to the task form to enter an answer);
- Every similar consecutive aggregated action in a tool was merged, and thus repeated actions to the same tool were counted only once;
- 2-gram frequencies were summed across all tools; each 2-gram (i.e., 2 tools) was compared to the total frequency of all 2-grams and converted to a ratio.

The result from this analysis appears in Figures 3 to 5, one figure for each of the tasks. The charts in these figures illustrate the direction of the tool use. The width of the lines indicates magnitude of that use; magnitude was measured by ratios and only ratios of more that 3% are included. The arrows show order of tool use within a 2-gram and the greyscale colours distinguish among the 2-grams.

For Task A (the selection of an appropriate video from a set of three), the pattern was not consistent. Task-Video and Slides-Task which represent the movement from completing the task response form to the ePresence interface seem evident. But many tools were used in many directions. For Task B (the response to specific questions), a distinctive pattern was apparent. Because the slides and toc factored significantly in responding to questions, participants often navigated between the two tools and the task answer form. However, the magnitude of these actions was relatively weak. Interestingly, the search tool is more frequently used in Task B than in Task A; participants were more inclined to use it when looking for specific pieces of information. For Task C (the writing of a gist), the pattern of user behaviour was much more distinctive, as participants used in a consistent fashion the slides and toc. Thus for Task A and B, the pattern of 2-gram usage was not distinctive, while for Task C a standard pattern emerged.

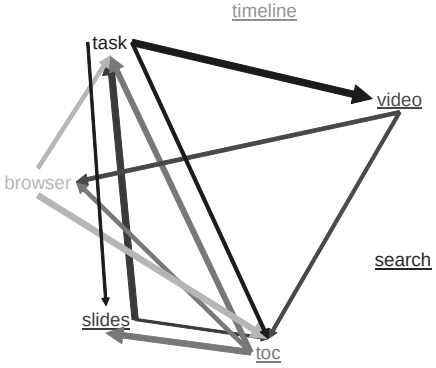

Figure 3. 2-Gram tool usage for Task A – Selection.

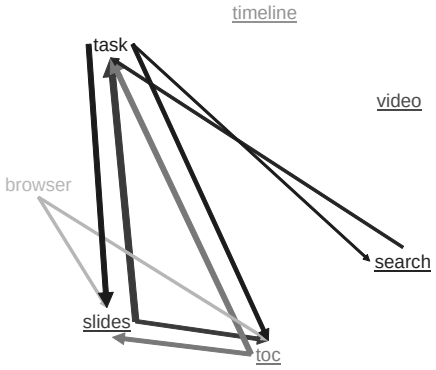

Figure 4. 2-Gram tool usage for Task B – Specific Qs.

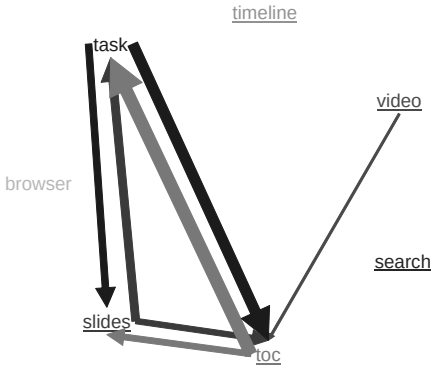

Figure 5. 2-Gram tool usage for Task C – Gist.

Participant's Tool Decision-Making

Our objective metrics show that participants approached tasks in different ways. As described under procedures, we re-played the screen capture video after each task was completed, and conducted a semi-structured interview to understand how and why participants made their decisions. This data helped us to understand the effects noted above.

Task A required a decision on the appropriateness of a webcast to a topic, a task analogous to selecting from a search results list, with the exception that there was no

abstract or textual transcript that summarized it. Participants claimed that this task was easier because "I did not have to find specific information." The complexity lay in extracting the necessary information on which to form an opinion. The slides became the tool of choice, as participant did a very fast skim of the slides to gain an understanding of the video. The toc was used like a large text summary. While the toc could have been a useful entry point into the video, participants tended not to click on a title. Given their time constraints, they did not want to waste time waiting for the video to load for something that might not be useful. When used, the toc helped to identify the introduction, overview or conclusion that a presenter normally provides.

Task B required accurate answers to specific questions, and the answers could not be found in the slides or in the toc, although both provided valuable clues. As one would surmise, participants tried and were somewhat dismayed about the functionality of the search tool. Search is not an especially strong feature of ePresence, for much the same reasons that it is a complex and problematic function in all video-based systems: voice recognition software is not yet robust enough to deal with content from multiple speakers. The search button only searches the text present in the toc and the slides. But participants had high expectations concerning how it should work. Thus with no access to lecture transcripts, participants were particularly ingenious about their use of the other tools. Their main strategy was to use the textual cues available in the toc and slides to find clues about where to "jump to" in the text.

Task C required the writing of a gist of the webcast. Participants had 15 minutes to understand the 45 minute video. They could not listen to the whole video, and had to make judicious choices concerning where to spend their time. Thus, they selected parts from the introduction, the conclusion, and discussion segments. The challenge was in identifying exactly where those segments were likely to be located. Participants tried the toc and the slides, and as it is clear from the two sets of data already presented concerning this task, they preferred the toc – "the agenda of the slides" – to the timeline in making those choices, as selecting a title from the toc 'fast-forwarded' the video to that point in the video. Slides were scanned to get a sense of the talk, as they were perceived as faster than listening to the audio.

In summary, participants' strategy was to make the system work for them, regardless of whether the best tool was available or functioned as expected. This was noticeable in Task A and B, but less so for Task C when the process used by participants worked for the assigned task.

CONCLUSION

Our study examined how people interact with a webcast to do classic information tasks. We found that certain types of tasks need certain types of tools, and until voice recognition research matches that done in other media, text is still king. In the case of our study, simply a text abstract or transcript would have simplified each of the tasks. However this is but a temporary solution; as participants pointed out, the video provides non-verbal clues about the confidence and abilities of the speaker; the video is not just an uninformative 'talking head.' By understanding user strategies, whether they are efficient or not, we have acquired a better sense of user needs in accessing webcasts, and a better sense of the types of tools that are needed.

ACKNOWLEDGEMENTS

Research funded by Heritage Canada and Natural Sciences and Engineering Research Council of Canada (NSERC) grants to the second author.

REFERENCES

1. Baecker, R.M. A Principled Design for Scalable Internet Visual Communications with Rich Media, Interactivity, and Structured Archives. In *Proc. CASCON 2003*, (2003), 83-96.
2. Geisler, G., et al. Video Browsing Interfaces for the Open Video Project. In *Proc. CHI'02*, ACM Press (2002), 514-515.
3. He, L., et al. Auto-summarization of Audio-video Presentations. In *Proc. Multimedia '99*, ACM Press (1999), 489-498.
4. Hürst, W, et al. The 'Authoring on the Fly' System for Automatic Presentation Recording. In *Proc. of CHI'01*, ACM Press (2001), 5-6. [Extended abstract]
5. Komlodi, A., and Marchionini, G. Key Frame Preview Techniques for Video Browsing. In *Proc. DL'98*, ACM Press (1998), 118-125.
6. Lee, H., et al. Implementation and Analysis of Several Keyframe-based Browsing Interfaces to Digital Video. In *Proc. ECDL'00*, (2000), 206-218.
7. Li, F., et al. Browsing Digital Video. *CHI Letters: Human Factors in Computing Systems 2*, 1 (2000), 169-176.
8. Rowe, L.A., Harley, D., and Pletcher P. BIBS: a lecture webcasting system. (2001) http://bmrc.berkeley.edu/ research/ publications/2001/160/bibs-report.pdf
9. Smeaton, A.F. Challenges for Content-based Navigation of Digital Video in the Fishlar Digital Library. In *Image and Video Retrieval*. Springer-Verlaged Heidelberg, London, UK, 2000, 215-224.
10. Toms, E.G., Freund, L. and Li, C. WiIRE: a Web Interactive Information Retrieval Experimentation system Prototype. *Information Processing & Management*, 40 (2004), 655-675.
11. Webopedia. http://isp.webopedia.com/TERM/W/ Webcast.htm
12. Whittaker, S., et al. SCAN: Designing and Evaluating User Interfaces to Support Retrieval from Speech Archives. In *Proc. SIGIR'99*, ACM (1999), 26-33.