Annotations for Streaming Video on the Web

David Bargeron, Anoop Gupta, Jonathan Grudin, and Elizabeth Sanocki

Microsoft Research, Redmond WA 98052

{davemb, anoop, jgrudin, a-elisan}@microsoft.com

ABSTRACT

Streaming video on the World Wide Web is being widely deployed, and workplace training and distance education are two key applications. The ability to annotate video presentations on the Web can add significant value by enabling "in context" note-taking and asynchronous collaboration through annotation sharing. We present preliminary data on the use of MRAS, a prototype collaborative video annotation system, for personal note-taking and for sharing notes.

Keywords

Video Annotation, Asynchronous Collaboration

INTRODUCTION

The use of streaming video on the Web for workplace training and distance learning has generated a lot of discussion and ideas lately. The ability to view content on-demand anytime and anywhere could expand education from a primarily "live" activity to include more flexible, asynchronous interaction [4]. However, key parts of the "live" classroom experience are missing from on-demand video environments, such as instructor responses to questions and comments.

A Web-based system for text or audio annotation of multimedia web content can alleviate shortcomings of asynchronous educational environments. It can allow students to record questions and comments as they watch Web-based lecture videos. It can allow them to share annotations with each other and instructors, either via the system or via email. It can enable them to watch their own and others' annotations scroll by as the video plays. It can support using annotations as a table-of-contents to allow jumping to relevant portions of the video. And finally, such a system can allow students to organize their annotations for personal and public use. It is in this spirit that we designed and built MRAS, our prototype video annotation system.

In this paper we compare the use of MRAS to the use of pen and paper for taking notes. We also look at how users build on each others' comments and questions asynchronously using MRAS, simulating "live" classroom discussion. We begin with a brief description of MRAS itself.

SYSTEM

MRAS is a Web-based client/server application which supports these features. The MRAS client is an ActiveX control that can be embedded in a web browser or HTML page. The user interface is displayed in Figure 1. The client communicates via standard HTTP with the MRAS server, which is itself a Microsoft IIS web server plug-in. Annotations are stored in a Microsoft SQL 7.0 relational database, completely apart from video content storage. For more details see [1].

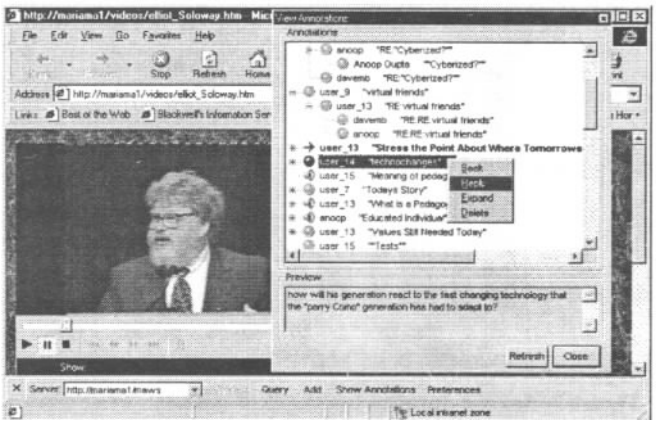

Figure 1: MRAS client user interface. The annotation view window is shown overlapping the Internet Explorer browser window. A user can skip to the location in the video (left) where an annotation was created. A tracking feature allows users to see the closest annotation to the current video time offset. Users can reply to previously created annotations, leading to threaded discussions.

PERSONAL NOTES USAGE STUDY

Our first study of annotation creation on streaming video explored the use of MRAS for taking personal notes.

Method

Six people were asked to prepare for a class discussion by watching and taking notes on the video recording of a lecture. Three participants took handwritten notes during the first half of the video and switched to MRAS for the 2nd half (Paper-first condition), and the other three did the opposite (MRAS-first condition).

Results

Taking notes using MRAS took participants 32% longer than on paper (this difference was significant at probability $p=.01$). This is mostly because MRAS automatically pauses the video when adding a new annotation (to avoid audio interference), whereas no such restriction exists for pen and paper. Despite the delay, all participants reported that the benefits of taking notes with MRAS outweighed the costs.

Another surprising result was that *all* of the MRAS-first participants paused the video while taking notes on paper; while *none* of the paper-first participants paused while using pen and paper. This suggests that people modified their note-taking styles as a result of using MRAS.

All six participants also indicated that their MRAS notes would be more valuable in the future (for example, for studying for an exam). MRAS annotations are automatically "contextualized" or linked to a video segment, which makes them easy to correlate with the video, and automatic correlation is supported in MRAS with "seek" and "tracking" features. Participants' handwritten notes contained little information that would help in correlating them with the relevant part of the video.

All six participants expressed preference for MRAS over pen and paper for taking personal notes, emphasizing organization, readability, and contextualization. This despite the fact that taking notes with MRAS took longer.

SHARED NOTES STUDY

The second study sought to assess the benefit of sharing notes. With a Web-based system such as MRAS, users can share notes with anyone else on the Web. Users can group and control access to annotations in ways that are impossible with paper-based notes.

Method

The same video was used in the "Shared Notes" study as was used in the "Personal Notes" study. However in the "Shared Notes" study there were 18 participants who used MRAS to take notes on the entire lecture video, adding text (typed) or audio (spoken) comments and questions to shared sets of notes.

Each participant took part in one of three conditions: Text-Only (participants could only add text annotations), Audio-Only, or Text-and-Audio. For each condition, six participants in sequence added notes to a shared annotation set. Each subsequent participant was encouraged to review and respond to annotations created by previous participants, as well as add new annotations. The first participant in each condition started with the same set of "seed" annotations.

Results

The number of annotations per participant (across all conditions) was not significantly different for adding new annotations ($p=0.45$), replying to existing ones ($p=0.35$), or for the combination of the two ($p=0.37$).

Four out of six participants using Text and Audio did, however, favor text over audio for creating new annotations and replies ($p=0.06$). These participants were also much more likely to reply to text annotations ($p=0.01$). Finally, participants generally felt it took more effort to listen to audio than to read text. Medium therefore had a subtle quantitative effect on the creation of annotations. Anticipating some of this (see [3]), we provided a text summary line for all annotations, regardless of media type.

We expected the number of new annotations being created to drop off over time, since we thought later participants would run out of new things to add. However, the total number of annotations added in sequence across the three conditions increased significantly (Pearson $r = 0.49$, $p = 0.04$), due to a significant increase in replies (Pearson $r = 0.52$, $p = 0.02$).

We hypothesized that inaccurately positioned annotations would be confusing to later users, since an annotation's position determines how it is contextualized by the video content. We found that people mostly positioned their annotations 10 to 15 seconds after the relevant part of the video (participants generally pressed the "Add" button a few seconds after formulating a response to something in the video). Contrary to our expectations, however, later viewers were not confused or distracted by the inaccuracy. When they watched the video from beginning to end, most found it natural to see annotations display in the preview window a few seconds after the relevant part in the video.

Participants generally liked using MRAS for taking shared notes on streaming video over the Web. The majority of participants reported that the annotations they saw from previous users were useful, a finding previously reported for annotation of text [2]. The vast majority also found MRAS to be an effective aid in preparing for a class discussion. 14 of 18 thought they made more comments than they would have in a live lecture. They were watching a video that could be paused and replayed, of course, but their response indicated that asynchronous Web-based support for annotation could increase class participation.

RELATED WORK

Prior work on annotation systems is reviewed in [1]. Such systems have been used to support guided pedagogy, help capture all aspects of live classrooms, provide feedback, and generally have been restricted to text as a target.

CONCLUDING REMARKS

The two studies of video annotation using MRAS described here indicate that the ability to annotate video over the Web can be a powerful means for supporting "in-context" note-taking and asynchronous collaboration. We are following up on these studies with experimental deployments in real-world classes and work situations.

REFERENCES

1. Bargeron, D., Gupta, A., Sanocki, E. & Grudin, J. Annotations for streaming video on the Web: System design and usage studies. http://www.research.microsoft.com/pubs/
2. Marshall, C.C. Toward an ecology of hypertext annotation. Proc. HyperText '98, 40-48. ACM, 1998.
3. Neuwirth, C.M., et al. Distributed collaborative writing: A comparison of spoken and written modalities for reviewing and revising documents. Proc. CHI '94, 51-57. ACM, 1994.
4. Stanford Online: Masters in Electrical Engineering. http://scpd.stanford.edu/cee/telecom/onlinedegree.html