Video Annotation in a Learning Environment

Xin Fu

School of Information and Library Science, University of North Carolina at Chapel Hill, CB# 3360, 100 Manning Hall, Chapel Hill NC, 27599-3360
 fu@email.unc.edu

John C. Schaefer

School of Information and Library Science, University of North Carolina at Chapel Hill, CB# 3360, 100 Manning Hall, Chapel Hill NC, 27599-3360 jschaefe@email.unc.edu

Gary Marchionini

School of Information and Library Science, University of North Carolina at Chapel Hill, CB# 3360, 100 Manning Hall, Chapel Hill NC, 27599-3360 march@ils.unc.edu

Xiangming Mu

School of Information Studies, University of Wisconsin at Milwaukee, Milwaukee, WI 53211 mux@uwm.edu

Numerous studies have demonstrated that annotation is an important part of human reading behavior in both printed and electronic environments. Annotation in the electronic environment requires special support due to limited media affordances. We have witnessed continuous improvement of annotation functions in some electronic reading environments, such as text documents in Microsoft Word or Adobe Acrobat and images in Flickr. However, comparatively little research has been conducted to understand people's needs for making annotations when they watch videos, let alone work to develop tools to support their needs. With the increasing use of videos in many aspects of our lives, from professional activities to personal entertainment, by not only specialists but also general consumers, there is need for more efforts on designing annotation facilities for video navigation and manipulation devices. This study focuses on video annotation in a learning environment. We studied how people in a teaching

assistant training class annotated videotaped instructional presentations. We attempted to understand the value of annotation in achieving their learning objectives and how video annotation functions helped in supporting their tasks. The results of this study provide implications for video annotation system design.

Introduction

Numerous studies have demonstrated that annotation is an important part of human reading behavior, both in the analog environment (O'Hara and Sellen, 1997; Marshall, 1997, 1998) and in the digital environment (Fu et al., 2005). Unlike the ease of scribbling when one reads a printed book, taking notes and making comments in an electronic reading environment can take inordinate amounts of time and effort, and be rather frustrating. Media affordance and the constraints of input devices require developers to include special techniques that provide readers of electronic documents with note-taking options similar to those they have access to in the print world. At the same time, we are interacting with a rapidly growing amount of digital information of various genres. While the computer applications for annotating some of these genres, such as Word and PDF documents, have become quite established among common users, comparatively little research has been done on video annotation. Video as a data type has only recently moved to digital form. The increasing availability and ever-decreasing cost of digital video capture equipment has resulted in video creation moving beyond the realm of specialists such as filmmakers and TV producers into the broad consumer market. Unlike text or still images, however, video sets the pace at which viewers experience it, presenting unique interaction and visualization challenges given its nature as an object existing in space and time (Ramos and Balakrishnan, 2003).

In this paper, we present the results of a study that explored and described the annotation behavior of a particular group of video users involved in a teaching assistants (TA) preparatory program. Specifically, the study focused on determining the annotation behaviors of the TAs when they interacted with video materials in a naturalistic setting; and through this, first understand their needs for making annotations and the values of the annotations to the annotators and the recipients of annotations, and second assess the value that the annotation functions may add in supporting tasks in this specific setting.

Background

In this section, we review technology and literature related to the use of video annotation. This background sets the stage for our study and frames the context for our design and

procedures.

Video materials have been created in two different formats: analog (film, VHS tapes) and digital (MPEG files, DVD). In either format, a specific device, be it a machine (VCR, film projector) or computer software (Windows Media Player), is needed to play the video. This constitutes the primary environment for video navigation and manipulation, similar to Web browsers for reading a HTML document or Acrobat Reader for displaying a PDF document. Thus, the development of video annotation facilities must be considered together with the embedding video navigation and manipulation tools.

There have been efforts to build annotation facilities into the video navigation and manipulation tools in both analog and digital environments. In the analog environment, Harrison and Baecker (1992) conducted task analysis for multiple users working with videos and noted that they worked in one of two ways: annotation and detailed analysis. With annotation, users "attempt to capture data in real time, in highly personalized and abbreviated ways" and the task is "characterized by high cognitive and attentional demands" (p. 158). Detailed analysis typically occurs after the real-time annotation and does not have the same constraints. Users may make many passes over a given segment of tape in order to capture more detailed information. Based on the task analysis, they derived a set of user requirements which support both the annotation and the detailed analysis process. The requirements were grouped into four categories: coding the data, analyzing and interpreting the data, user interface and device control, and displaying the data. They also laid out the specifications for designing video annotation tools and illustrated how they were implemented in a specific tool called VANNA.

In the digital environment, Ramos and Balakrishnan (2003) described a prototype system called LEAN for controlling and annotating digital videos. The system explored a variety of interaction and visualization techniques for fluid navigation, segmentation, linking, and annotation of digital videos. In particular, the tool used pen-based technology and supported frame or segment-based annotations. The authors also elaborated on how annotations referring to objects that are temporal in nature, such as video, might be thought of as links, and fluidly constructed, visualized and navigated.

The notion of building links is also explored in another study that designed an application called the Interactive Shared Educational Environment (ISEE) to allow video annotation/discussion via a chat panel in sync with the video (Mu et al., 2003). A customized version of this application is used in this study. Details of the tool will be presented later in the paper.

All these studies suggest that video annotations follow similar paradigms as annotations of other data types. A video annotation tool, like other types of annotation tools, must

provide a set of three functions: selecting the segment of the annotated information object (like the anchor in hypertext), adding the notes and creating the link, although the way each function is implemented in specific systems differs considerably.

Unlike the flourish of annotation functions in some of the electronic reading facilities such as Microsoft Word and PDF Reader, and even in image viewers such as Flickr, mainstream video players like Windows Media Player and Real Player have little support for annotation. This gap triggers our interest in gaining a better understanding of people's annotation needs when they work with videos. In previous literature (e.g., Schilit et al., 1999; Fu et al., 2005), it is observed that users often resort to workarounds such as printing out the documents and making annotations on the printed version when they work with digital documents such as text files or Web pages. Although it is impractical to print out a video to annotate it, do people also tend to take notes on paper when they watch videos? If so, what will be the additional benefits of integrating annotation functions into the video player? Will people feel differently if they take notes with a video annotation tool? Finally, are people's annotation needs and behaviors different or similar in video browsing versus text reading and how does this affect design?

Methodology

We designed a naturalistic study to observe participants' annotation behavior and used a number of data collection instruments including interview, questionnaire, and focus group. Our design provided us with a rich understanding of the annotation behaviors of our participants in a specific setting. Although we can not generalize the results to the universe of digital annotation, we believe that they provide unique insights into the annotation behaviors of our participants and may lead to the creation of more effective annotation tools for similar uses. In this section, we describe the participants, the study procedure, and the video annotation system that was employed in the study.

Participants

The participants of this study are the members of a graduate level course offered in the authors' institution. This course, as a part of the Preparing International Teaching Assistants Program, is designed to help international teaching assistants prepare for their role as classroom or lab instructors by improving their English pronunciation, cross-cultural communication and teaching skills and increasing their communicative competence in the American classroom. Seventeen members of the Fall 2005 class from three sections (the instructor and 16 students) volunteered to participate in the study. Each section meets twice a week for 50 minutes.

The curriculum of the course involves four videotaped presentations per student, designed

to simulate classroom lectures on their respective subject area (e.g., chemistry or statistics). Each presentation is focused on a specific theme (e.g., self-introduction or defining a term) and lasts approximately 10 minutes. The format includes additional time for mock audience questioning. With seven to eleven students in each section, it usually takes three class meetings to videotape all of the presentations. The next three class meetings are review sessions in which the instructor replays each presentation to the class. Sometimes, the instructor stops the video replay at midpoint to make comments. After watching each presentation, class members receive a peer feedback form and spend a few minutes to write down their feedback for the presenter (including themselves) guided by structured questions related to the theme. Then, the class holds an open critique session on that presentation to allow further clarification and discussion. At the end of each review class, the instructor collects the peer feedback forms and distributes them to the appropriate person (i.e., the individual who gave the presentation).

Procedure

The study took place around the middle of the semester, between the second and the third presentation arranged in the class schedule. After obtaining IRB approval, we announced the study in all class sections and distributed the study consent forms. Class members interested in participating in the study were asked to bring the signed consent form to a 10-15 minute pre-study interview.

In the interview, we collected the participants' opinions about the process of making and receiving comments using paper forms in their first two presentations. We focused on the usefulness of the form in eliciting comments and reviewing the comments. In particular, we paid attention to how the structure of the form facilitated the review process. All 17 participants participated in the pre-study interview.

In the week following the pre-study interview, students gave their presentations, which were recorded and burned into a CD that contained the video navigation and annotation tool (i.e., ISEE, described below) as well as the presentations of all the students in the same section. Each class member received a CD of their own section, watched and annotated the videos outside of class., and submitted the files to the study team.

After collating the files, we reviewed them to determine how and why people made annotation. We also considered the content of the annotations and how it differed from the content of the paper-based peer review forms. The saved comments were collected and forwarded to relevant presenters for their own review.

After completing the review process, each participant completed a questionnaire to document their experiences both as a commenter and as a recipient of comments, and to

compare video annotation to their experiences using paper feedback forms. Fourteen of the participants completed the questionnaire.

Finally, we observed a class session in which participants shared feedback based on the video annotation process, interviewed the instructor, and held two focus groups that involved a total of 14 student participants. Both the interview and the focus groups took place within one week of the use of the video annotation tool and personal review/reflection process. The results tended to relate to ISEE's functionality, its use as an alternative to a paper-based process, and the value of video annotation in the learning environment. The study procedure is summarized in Table 1.

From the instructional point of view, the process of reviewing the presentations is changed with the new tool. The differences are highlighted in Table 2.

Table 1. Study procedure

Investigators	Participants	
Submit IRB		
Enroll participants		
	Sign consent forms	
Conduct pre-study interview	Participate in pre-study interview	
	Make and record presentations	
Make CDs		
Distribute CDs and questionnaires	Obtain CDs and questionnaires	
	Review presentations and make comments	
	Submit comments	
Redistribute comments		
	Review feedback	
	Fill out questionnaire	
Observe discussion	Discuss annotation in class	
Conduct interview/focus group	Participate in interview/focus group	

Table 2. Presentation review process with and without the video annotation tool

Step	Without ISEE	With ISEE
Record presentation	Takes 3 class meetings	Takes 3 class meetings

Receive feedback form	Before reviewing presentation	Before making presentation
Replay presentation	In class, takes 3 class meetings	Outside class, takes varied amount of time depending on student motivation and section size
Instructor comments	Before the whole class	Sent to individual student
Open critique	3-5 min each	About 10 min in total
Give feedback	Fill out structured paper forms in class (3-5 min each)	Type comments on ISEE with printed guidelines (unlimited time)
Review feedback	Read paper forms w/o video	Read comment files w/ videos

Video Annotation System

ISEE is a multimedia distance learning and collaboration tool that is specifically designed for low bandwidth network users (Mu, Marchionini, & Pattee, 2003). In addition to video navigation functions, ISEE includes a video annotation function that freezes the video when the user starts to enter a note in the input box. It continues when the user hits enter or presses the "submit" button. The user can also apply time stamps that link the notes to video segments. A full version of ISEE contains a video player, an interactive chat room, a built-in web browser, and a story board. We reconfigured the tool for this study by excluding the web browser and story board components to help the participants focus on making the notes and minimize the distraction of other functions. The interface is displayed in Figure 1.
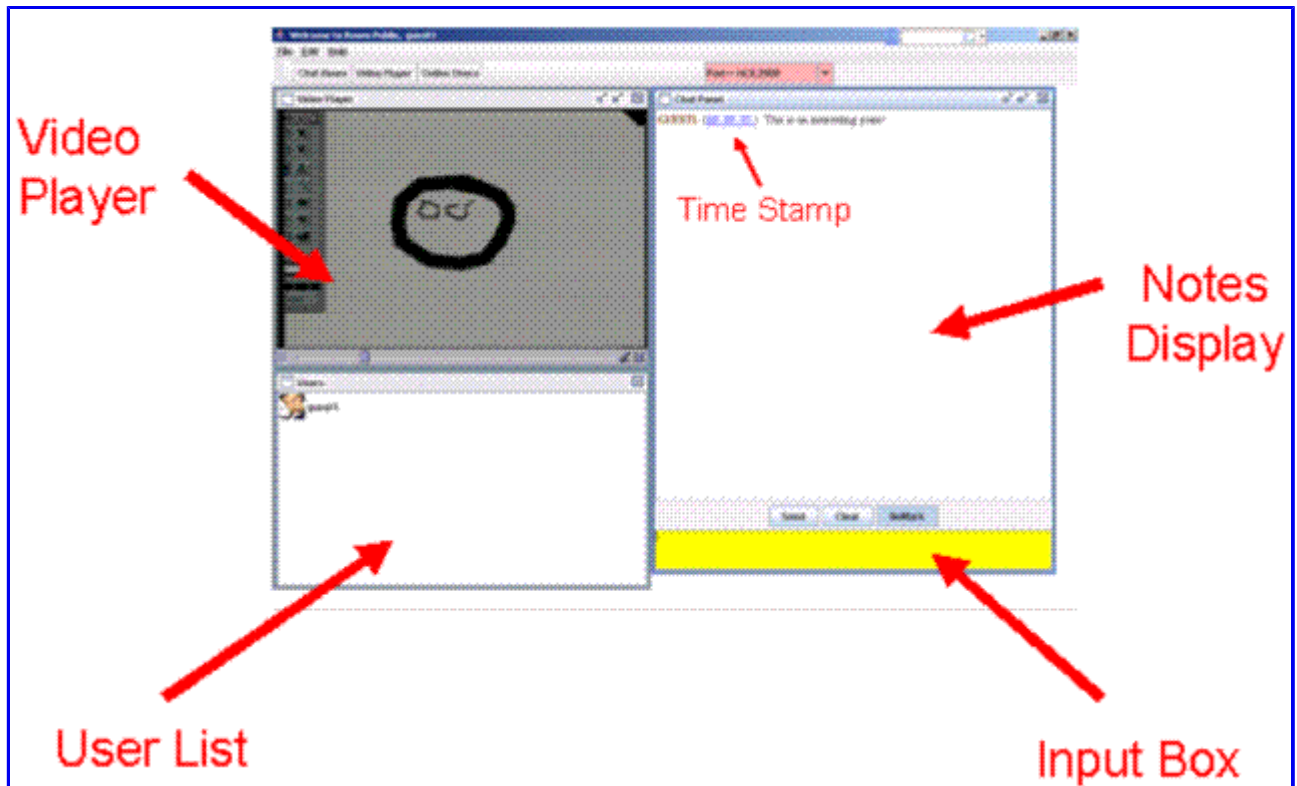
Figure 1. ISEE interface

Results and Discussions

This section presents the findings of the study and discusses their implications. We first describe how the participants take notes with and without ISEE. Then, we report the findings on the values that annotations and the annotation process bring to people. Next, we analyze how annotation tools facilitate the making and use of annotation. We also summarize the functions that the participants expressed a desire for us to incorporate into future video annotation tools. Based on the analysis, we put forward some design specifications.

Annotations on Paper

As previously noted, students currently view videos in class and make comments on highly structured forms customized for particular course objectives (e.g., introducing a syllabus, using a visual, defining a term). Students use the one to two page forms to provide constructive feedback to their peers with the intent of improving their presentations. The forms include both open and closed questions (often with multiple choices) and check lists. Students complete the forms in class while watching the presentations on a LCD projector. In some cases, the instructor stops the presentation to identify and discuss good or bad practices.

A review of a small set of completed forms suggests that few students complete the entire

form, and fewer offer extensive comments. In most cases, students simply indicated yes or no or rephrased the question into a short answer.

Annotations in ISEE

All participants in this study used the annotation function to comment on their colleagues presentations. According to our results, participants annotated videos nine times on average. The length of the annotations ranged from one to two words up to a maximum of 113 words, with an average length of ten words. A review of the time stamps in the annotation files demonstrates that participants made annotations throughout the length of the video.

Typically, most participants started to annotate the video as soon as the presenter touched on subjects reflected in the feedback guidelines. For example, the first item on the guideline asked participants to answer the question "Did the speaker state the formal definition of the term?" When discernable, a majority of the participants' initial annotations correlate to this question. Other correlations to the guidelines appear throughout the majority of the participants' annotation files. The only significant variation occurs in cases where the annotator appears to be highly motivated to provide detailed, lengthy annotations. These individuals (approximately 30% of our population) tended to create annotations of a much more expansive nature.

The review of the annotation files also suggested that people used shorthand and other shortcuts to facilitate the annotation process. The use of these time-saving techniques tended to vary somewhat depending on the purpose of the annotation (i.e., peer review or self reflection). Several of the participants used emoticons and other symbols to convey information to or emphasize points for their colleagues during the peer review process. Examples include:

> *You are smiling :) that is great :)*
> *He tried to explain democracy concept by eminent scholars' comments=> Very Nice*

In addition, one participant used all caps to emphasize that a colleague provided the wrong explanation for a statistical concept. Another participant appeared to use a larger font size to emphasize an entire presentation. Finally, other participants relied on exclamation points and other punctuation to reinforce comments. Figure 2 exemplifies an average annotation in our set.
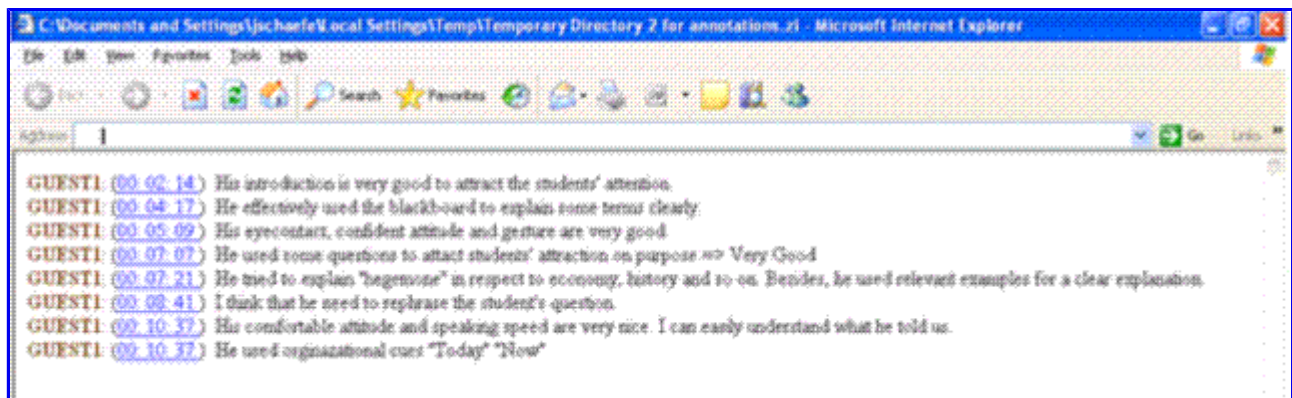
Figure 2. An average annotation file

When annotating video for self reflection, participants used more shorthand and phrases that required personal understanding to decipher. The analysis also suggests that individuals created more annotations during the self reflection process (see Figure 3 for an example).
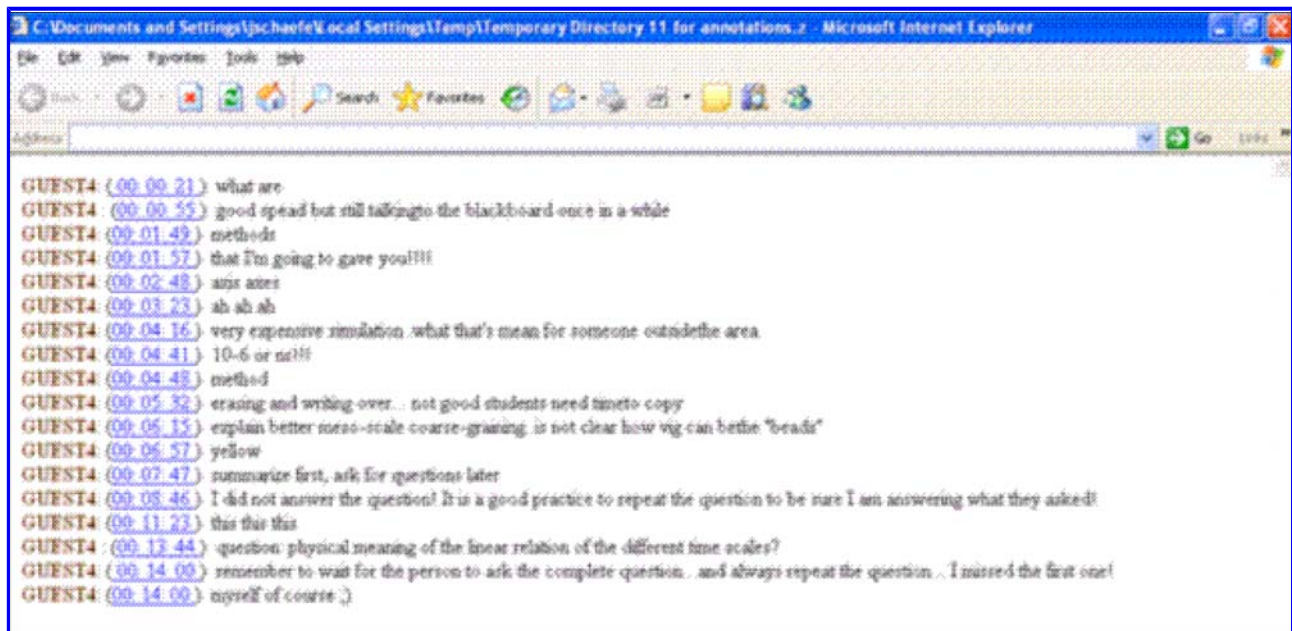

Figure 3. A self-reflection annotation file

Use of Time Stamps

The majority of participants used the time stamp function when annotating video. A review of the annotations suggests that most of the participants used time stamps to link annotations to specific events that they observed in the video. A look at a few typical comments (extracted from multiple files minus the time stamp) demonstrates how participants used the function to link their comments to specific portions of the video.

Too much time talking to the blackboard. You should face your students and do eye contact

Very good! She is encouraging ... students to ask ... questions

Well prepared visual, good job with explaining the graph in a simple manner

It's better to read something completely and clearly rather than only reading part of [it]

Some participants appeared to have difficulty using this feature. For example, some of the annotations appear to refer to specific events in the video yet the annotations lacked a time stamp. Other annotations appear to have been general in nature but included a time stamp that linked them to a specific portion of the video. In a few cases, it is possible to determine that a person deliberately deactivated the time stamp feature to create a more general type of comment. As Figure 4 shows, this participant deactivated the time stamp function to offer a general comment on the presenter's performance and to create a question for later reflection.
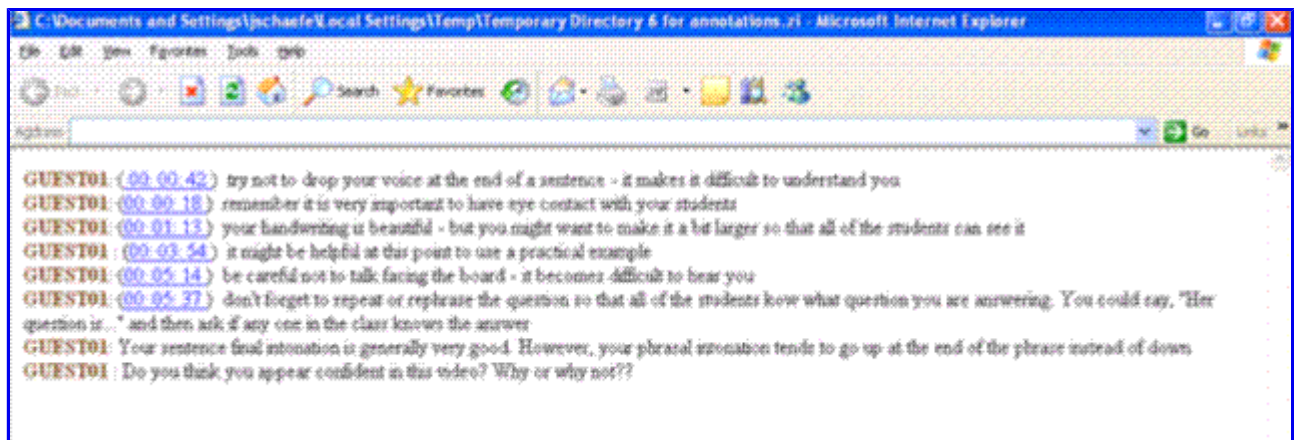
Figure 4. Alternation in time stamp use

A Look at Annotation Content

The length and quantity of the comments received in the annotation files suggest that people provided more feedback than what they typically provided on the peer review forms. As common sense would suggest, the results also indicate that the longer the annotation, the more substantive the information. As shown in Figure 5, the longer annotations demonstrate a detailed understanding of the topic at hand as well as useful suggestions on how to improve delivery. It also shows the importance of linking specific annotations to video frames because in this case the annotator evidently deactivated the function inadvertently halfway through the presentation.

After analyzing the data, we grouped the annotations into three categories based on length and general content.

One word. One participant used one word annotations to provide information on how to pronounce words. It is interesting to note that the annotator divided the word into syllables

to facilitate correct pronunciation (e.g., 're-pre-'sents, 'mito-'chon-dria, dif-fer-'entiate). Other participants frequently used one word annotations to reinforce good practices or simply to provide supportive comments.

Two to five words. Annotations of this length typically provided observational type feedback such as nice smile, speak slower, or good explanation. One participant used the term check comprehension on two occasions. Many participants created short phrases to reinforce good practices (e.g., good eye contact, Good facial expression and gesture!) or to point out areas for improvement (e.g., she can be more enthusiastic!). In some cases, participants simply repeated or rephrased some of the information in the guidelines at applicable points in the video.
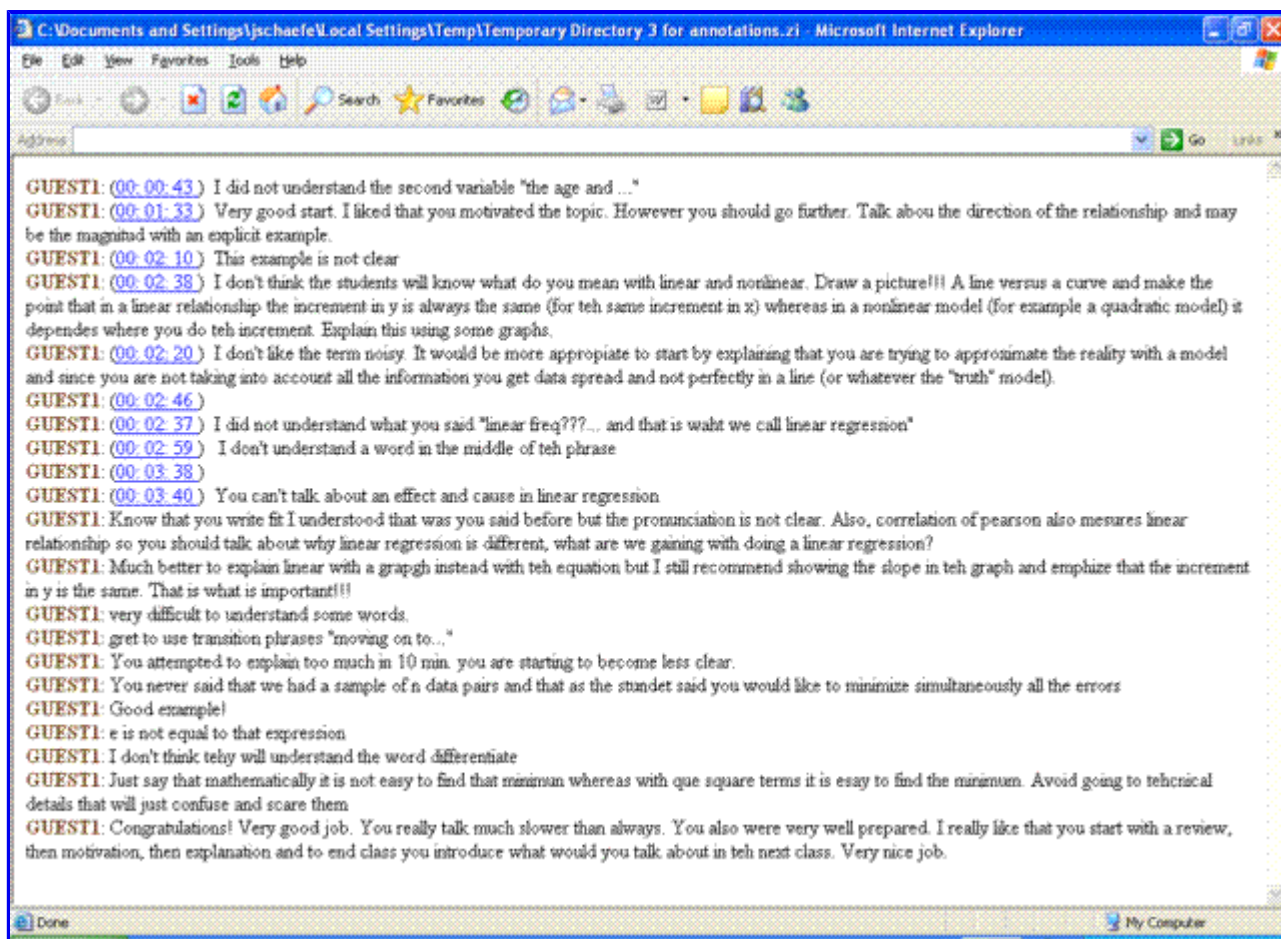
Figure 5. A longer annotation file

More than five words. Annotations longer than five words chiefly consisted of two types: 1) observations on teaching performance and suggestions for improvement; or 2) specific comments or criticisms about the delivery of domain specific information. In many cases, the observations on teaching performance provided detailed information that participants commonly noted was missing in the structured forms. Instead of just reflecting what people mentioned in class, which participants stated routinely occurred in the old,

form-based process, the video annotations offered detailed suggestions on how to improve performance. After viewing one video, for example, a participant proposed that his colleague should take the following steps to improve his performance.

> *For your next presentation, this would be a great place to get the students involved with interactive questions - in other words, they could give you the answers and you would then know that they understood*

Several participants' responses also reflected detailed suggestions on how to improve performance and the delivery of domain-specific information. This type of annotation commonly contained abbreviations, terms, and variables that people from outside of the domain may have a hard time understanding. As the following example illustrates, the participant making the annotation was confident that the receiver of the comments would understand the content of the note.

> *My suggestion is for next presentation just talk either the discrete case or the continuous one. Also, you need to emphasize that sometimes is impossible to list all values for discrete rv (example poisson). Otherwise it seems that discrete cases are finite. For the continuous case, start by making them realize that the probability of x being a specific number is zero and hence we are interested in intervals.*

Value of Annotation

The values of annotation as perceived by the participants are mainly manifested in two aspects: the value of the comments to the recipients and the value of the annotation process to the annotators.

Value of the Comments for the Recipients

Most participants considered the comments that they received from the peer feedback forms more or less useful for planning future presentations. One individual stated that he reviewed the comments because "you cannot see yourself make presentations." From his perspective, comments from others provided useful information about how people interpreted his voice inflections, facial expressions, and hand movements. This information helped him to improve his teaching skills.

Participants felt the comments they received from ISEE useful, too. In addition to the values that comments on paper forms bring, the participants emphasized the benefit of receiving detailed comments that pinpoint a specific event, action, or element in the video. For example, one participant mentioned that he valued the feedback because it let him "know exactly what part ... to improve [based on] the time mark."

Compared with the classmates' comments, the instructor's comments were deemed much more helpful by the majority of participants. As a native English speaker, the participants believed that the instructor understood their presentations better and therefore could make longer and more detailed comments. (The perceived value of the instructor's comments is also most likely due to her expertise and years of classroom experience.) As a result, they usually spent more time reviewing the instructor's comments than those from their colleagues. We observed that one participant kept the instructor's feedback for future reference, but not the peer feedback. Figure 6 (overleaf) contains an example of a typical feedback form filled out by the instructor juxtaposed to a typical feedback form filled out by a peer student.

Despite the value of the comments in general, almost all the participants stressed the wish to receive more responsible comments from their classmates. Sometimes, the reviewers were too polite to give useful comments. In the paper based review method which involves an in-class open critique, the participants found their classmates merely repeating on the feedback form what the instructor had said about the presentations or what had been brought up in the open critique. Students frequently write few comments and often leave much of the peer feedback form blank (see Figure 6b for an example). In other cases, especially in paper based review, participants received many comments too general to be useful. For example, one participant noted that he received a comment from a feedback form saying "incorrect grammar" without the context necessary to identify and correct the specific mistake. Some participants found general comments like "lack of eye contact" or "voice too low" useful but most participants felt that general comments like "I did not understand you" were not useful unless the commenter included data to place the comments into context.

Figure 6. Typical peer review forms (Left: from instructor; Right: from student)

Value of the Annotation Process to the Annotators

In both the pre-study interview and the post-study interview/focus group, the participants mentioned the value of the annotation process to the annotators. A participant mentioned that the peer feedback forms helped him to think about the presentation before hearing it. The forms also provide him with the ability to think about his own presentations, and the opportunity to memorize what he must talk about. Another participant noted that the need to give comments and feedback made her more attentive during other people's presentations.

This view was echoed by the instructor. During her interview, she emphasized that the process of giving comments to colleagues was designed to offer students with an opportunity to internalize the points outlined in the feedback form and learn by observing other people's good and bad practices.

Comparison of Methods

A qualitative analysis of the comments on paper-based feedback form and the annotation files created in ISEE suggests that the later provided more detailed and useful feedback than the paper forms. A common complaint voiced by participants in the pre-study interview was that the short and general responses they received from colleagues were typically of little value. The video annotation tool provided participants with the ability to specifically link annotations to precise moments in time. This connection created a context in which participants could provide more valuable feedback through a combination of image and text.

Quantitatively, we also noted that the participants provided longer and more detailed comments through ISEE than they normally did on the paper form. As noted above, the longer annotations demonstrated a detailed understanding of the topic at hand as well as useful suggestions on how to improve delivery.

Usefulness of Annotation Tools

In this section, we summarize the findings on the usefulness of annotation tools and what additional values they bring to the annotation process. We report the participants' likes and dislikes for the two types of tools that we studied: the peer feedback forms and ISEE. For each tool, we consider its usefulness for the reviewers and for the recipients of the comments.

Peer Feedback Forms: From the reviewer's point of view, the participants expressed mixed feelings towards the usefulness of the forms. The majority of participants stated that the structure of the form was useful because the questions highlighted what to look for during presentations. A participant said that the forms helped him be organized when he commented on the presentations. The specific questions guided him through the comment process. Some participants also noted that the structure of the form was helpful when they reviewed the feedback. A participant said he could easily pick up the comments addressing the same issue from the same place on the form, so that he could review the feedback aspect by aspect, instead of form by form.

However, other participants identified several problems with the forms. A common complaint was that the forms were too long and the participants did not have time to carefully respond to all the questions. Some participants criticized the form's structure, stating that the specificity of the questions limited their ability to provide unstructured feedback. One participant mentioned that some of the questions were so similar that they became repetitive (e.g., *Did the speaker show (a) concern, (b) interest, (c) openness, and (d) confidence?*). Some participants found the open-ended questions difficult to answer

due to the limited class time afforded for this task and the complexity of the questions, which posed challenges for the nonnative English-speaking students. For example, one participant mentioned that he did not know how he could illustrate the problem of "*lacking care about the student*" in simple sentences. Another participant said that it was much easier to evaluate and provide feedback using the check list items included in some of the forms instead of developing the criteria necessary to answer open-ended questions on his own. For example, with regard to a question, "*Did the speaker use any classroom terms/expressions? If so, which ones?*" the participant noted that it was not easy for reviewers to remember terms or expressions used in the presentations after a single review session.

Video Annotation Tool: We used questionnaire and interview/focus group to capture feedback from participants related to the usefulness of ISEE in the annotation process. In general, most participants identified the time stamps and video control functions as the things that they most liked about the tool, both as reviewers and recipients of comments.

As noted before, the majority of participants used the time stamp function when annotating videos to link annotations to specific events that they observed in the video. This saved the trouble of having to describe the event itself in words; instead, the reviewer could focus on comments and suggestions for improvement. One participant noted that, because of the time stamp function, it is "*more convenient to write notes [with ISEE]*".

The time stamps were also helpful when the participants received and reviewed the comments. Linking the comments to the places that they referred to helped the reviewers make better sense out of the comments and understand the comments within the context. Not surprisingly, the participants noted that such a function was most helpful when a comment pointed to a specific place, and less so if it was a general comment about the whole presentation.

Another feature that participants found useful about ISEE was the function to control the viewing process. Individuals can use the tool to select a specific video from an options list, decide on how much of the video that they would like to review, select a particular scene or event to annotate with the video freezing automatically, or even resize the window to fit their needs. Several participants also mentioned that they appreciated the ability to rearrange windows to make the video more prominent than the comments (to focus on the video) or to make the comments larger than the video (to review the previous work).

Another participant's response not only illustrates the benefits of being able to control the video and the annotation process from the user's perspective, but also reflects the sentiments of the other participants with regard to the value of video annotations.

It helps you to write comments as soon as you see something worth commenting. In the case of first watching the video and then reviewing it, as we have done [in class] before this, one tends to forget by the time the video ends, and so the feedback is incomplete in many respects.

In general, participants seemed satisfied with the use of ISEE to support the video annotation process. Participants viewed it as a tool that allows them to interact more intimately with each presentation, giving them the possibility of not only greater exploration, but introspection in that they consider their own work more closely. With almost every participant in the study, we heard comments that ISEE facilitated their interaction with the video presentations and improved their ability to provide meaningful comments. While paper forms allowed for more general observations, which help clear up more widespread problems, the annotation tool enabled participants to provide their colleagues with exhaustive feedback and outline more issues.

Some participants found ISEE less helpful or frustrating due to the technical difficulties that they experienced while using the tool. Typically complaints included:

Sometime it is not easy to log in and the program breaks down some time. I have to restart the program again.
Can't edit comments.
It always happens that whenever I try to save my comments, the ... windows disappear.
I lost all what I have done in the last 30 minutes because it [closed]!

Two other comments related to the use of a chat room and the difficulties that resulted in having multiple users of the room commenting on different videos. Our study anticipated this problem and created a procedural workaround to prevent it from impacting out study. The result does suggest the need to create an offline annotation function.

Another factor which limits the usefulness of annotations made through ISEE is the lack of structure. Before ISEE was introduced into the class, participants had been using highly structured forms to comment on teaching presentations. While somewhat restrictive, the structure of the form does result in a set of organized feedback. One participant noted the lack of structure in the comments sent to him by his colleagues. His quote below reflects an important difference between structured versus unstructured information.

No, there's a problem in this aspect, because the comments are very general, and not categorized into specific sections. It's like a mass of data that has all the information in it, but is so disorganized that it can't serve its purpose unless statisticians put it into

presentable format and remove whatever is extraneous.

Finally, at least one of the participants commented that a shortcoming of exchanging feedback through ISEE is the lack of conversation between participants that routinely occurred in the class-based review sessions:

> *[ISEE is] more efficient, since if we do the review session in class, nobody can guarantee that he remembers the whole presentation and make precise comments. But with the video in hand, we can replay it if we don't think we get enough, and then find more places to comment about. The problem is, if we discuss in class, then there's not only comment, but also response. Say, if somebody comments on my presentation, then I can reply, I can explain what I was thinking when making the presentation, and other people can also join us, discussing and extending to some more topics about the presentation. But with the note taking function, we can not do it at present, right?*

Suggestions for Future Improvements

In regard to a question about additional features, participants provided useful feedback to guide potential future modifications. In addition to general suggestions to improve the basic stability of the tool to address the issues noted above, participants asked for specific features to improve the annotation function. People would like to add color to emphasize notes, have access to a template or pick list that includes words or phrases frequently used in a review process (e.g., good comment, useful slide, nice explanation), and have the means to classify and organize annotations. Individuals would also like the ability to edit comments after submission. Some participants said they would like to see greater flexibility in time stamps to account for different levels of granularity. Instead of always linking to a specific time point, they would like to have different types of time stamps, some of which link to a point, some to a segment, some to a chapter (if such a structure exists in the video), and some to the entire video. One participant even raised the need of annotating a particular area in a key frame, which alluded to the topic of image annotation.

Summary of Results

Generally speaking, these results provide a glimpse at the types of annotations that participants of this study make in a digital environment. While not conclusive, our analysis suggests that both paper-based annotation and video annotation support the learning environment to a similar degree, but video annotation might offer a slightly improved environment due to its ability to provide a more detailed context for peer review through the simultaneous application of three components-i.e., digital video, timestamps, and

annotation content.

The results indicate that most of the participants annotated video at various points throughout the presentation. A majority of the participants closely followed the guidelines distributed in class to facilitate the review process. While the guide apparently triggered most annotations, some participants tended to select specific moments, gestures (or lack of), or words to motivate them to make an annotation. In many cases participants created their own emoticons and used other shortcuts to convey their feedback to their colleagues.

The results of this study highlight the important functions for video annotation tools in this specific setting. For example, the use of the time stamps demonstrates the importance of providing a means to link the annotation to the object of the annotation. A function of this type saves time and effort on the part of the annotator because he or she does not have to spend additional time creating a context for their comments. This appears to allow individuals to spend more time on developing the content of their annotations. The finding resonates with and provides an illustration for Ramos & Balakrishnan (2003)'s argument that the nature of video as an object existing in space and time presents unique interaction and visualization challenges.

The result also suggests that it is important for a video annotation system to provide support in the following areas. First, a video annotation system should provide control for video navigation, such as pausing, fast forwarding and rewinding, during the annotation process. Next, a video annotation system should allow users to edit and modify the annotations. In addition, it is important that video annotation systems provide functions to allow users to specify the granularity of the annotation. A particular annotation may refer to a single frame, a segment, or the entire video. Finally, a video annotation system should incorporate some design for customization. The customization functions will build user satisfaction.

Our results suggest that the tool could be improved to include features such as a pick-list for frequent comments or phrases connected to class requirements. The pick-list could be both domain and non domain specific. For example, in the case of peer review of teacher performance, the pick list could include short phrases that indicate when a participant was within acceptable parameters (e.g., *good explanation, informative visual, good organization of data*) and when they deviated from acceptable limits (e.g., *speak slower, confusing answer, lower your voice*).

In the end, the value and use of annotation depended primarily on the behavior and motivations of the individual participants. Our results clearly indicate that some participants spent a great deal of time reviewing and commenting on their colleague's

performance. A general comparison of the number of and length of annotations per file results in a natural separation of highly motivated reviewers from those individuals who perhaps approached this assignment in a more casual manner.

Conclusions and Future Directions

It is clear that the four parts of the study provide consistent evidences to inform the research questions. First, the results suggest that both annotators and the recipients of annotations benefit from the annotation process. The process assists annotators by keeping them attentive when they watch the video and helping them internalize the learning objectives. The process of annotating over other people's presentations deepens their understandings of guidelines that the instructor provided for preparing the presentations and helps improve the annotators' presentation skills as well. The annotations benefit the recipients in a more content related way. The annotations reinforce the presenters' strong points and build their confidence, make them aware of their weaknesses, and help them to set goals for future efforts. Moreover, the annotations represent a form of interaction among class members. The shared task of annotation supports the creation of a community focused on fostering improvements in all of its members. Some individuals clearly stand out as community leaders demonstrated by the quantity and quality of the feedback that they share with members.

Second, annotation functions facilitate the annotation process in different ways. As discussed in the above section, the time stamp function facilitates the annotation process by providing context and building connection for both reviewers and their subjects. The context and connection dramatically improve the review process by providing actionable feedback and directions. Structured paper review forms may provide some context but results suggest that they limit an individual's ability to provide meaningful, free-flowing comments.

In general, we found that participants prefer the possibility of working directly with the video, instead of resorting to alternative means, such as the paper form in this case. It is important for video annotation tools to account for the temporal nature of videos and provide flexible methods for users to specify anchors and building links. It is also important to integrate video navigation and manipulation functions into the annotation system.

We cannot conclude this paper without acknowledging some limitations to this study. Although we walked our participants through ISEE and gave them a detailed step by step instruction with a screen shot for each step, we did not provide an opportunity for them to gain some experience with ISEE before they started using it for their tasks due to the constraints of the project time frame.

As noted in the results section, we observed similar practices when people made comments on the videos in paper and electronic environments. We also observed that the quantity and quality of the annotations were different. However, our analyses were inconclusive to attribute this difference to any one of the several possible factors, such as the effect of the structure of the paper form, the naturalistic setting for watching the video and making annotations versus the in-class setting, and the effect of using an annotation tool instead of the paper form. Our results suggest that the practices and values of annotation in both environments differ greatly based on tasks and personal motivations.

We are starting another round of data collection to address some of the above limitations. For example, we plan to distribute a practice CD to the students so that they have an opportunity to practice with the tool prior to having to use it to review and annotate video. We will draft some short tasks that involve opening and closing a video, activating and deactivating the time stamp function, creating and saving annotations, and sending the file electronically. We hope that this change will allow students to familiarize themselves with the video annotation tool prior to using it to review and comment on their classmates' videos.

Finally, we plan to conduct a future study to examine the social aspect of annotation using the collaborative function of ISEE. We are interested in learning how people's behaviors change when they simultaneously watch and annotate video from different locations. It will also be interesting to study how the interaction in that collaborative annotation process differs from the interaction taking place in the in-class open critique.

Acknowledgements

References

Fu, X., et al. (2005)   Annotating the web: an exploratory study of web users needs for personal annotation tools   *Proceedings of the 68th Annual Meeting of the American Society for Information Science & Technology, 42. Charlotte, NC*

Harrison, B., & Baecker, R. (1992)   Designing video annotation and analysis systems   *Proceedings of the conference on Graphics interface '92, 157-166. Vancouver, BC*

Marshall, C. (1997)   Annotation: from paper books to the digital library   *Proceedings of*

the ACM Digital Libraries '97 Conference. Philadelphia, PA

Marshall, C. (1998)   Toward an ecology of hypertext annotation   *Proceedings of ACM Hypertext '98, 40-49. Pittsburgh, PA*

Mu, X., Marchionini, G., & Pattee, A. (2003)   The Interactive Shared Educational Environment: user interface, system architecture and field study   *Proceedings of the Annual Meeting of the American Society for Information Science and Technology* 40, 291-300

O'Hara, K., Sellen, A. (1997)   A comparison of reading paper and on-line documents   *Proceedings of CHI '97* 335-343. ACM Press

Ramos, G. & Balakrishnan, R. (2003)   Fluid interaction techniques for the control and annotation of digital video   *Proceedings of ACM UIST '03, 105-114. Vancouver, B.C.*

Schilit, B., et al. (1999)   As we may read: the reading appliance revolution   *IEEE Computer* 32(1), 65-73