

An Ink and Voice Annotation System for DENIM

ABSTRACT

I have implemented ink and voice annotations in DENIM, a sketch-based web site design and prototyping tool. I also performed an informal evaluation of the annotation feature with four participants, one of whom was a professional web designer. The evaluation focused on how people use ink and audio annotations to review existing designs of web sites. The participants in the experiment tended to use many more words in voice annotations than ink, and they used a more conversational tone with voice than with ink, which were usually short sentence fragments. Out of the four participants, one preferred ink, one preferred voice, and two had no preference for these specific tasks, although all four perceived a need for both types.

INTRODUCTION

DENIM [5] is a sketch-based prototyping tool for designing web sites. It allows designers to design web sites at multiple levels—site map, storyboard, and individual page level—and specify behavior by linking pages with arrows. Designers can then interact with their web site design through a Run mode.

Although DENIM is designed to help designers through the early stages of design, it currently lacks some key features that would enable DENIM to be fully integrated into a web designer's work practices. One of them is annotations. When designers sketch out their designs, they also annotate them with notes that help them remember design decisions and clarify and motivate their designs. When designers get together to discuss their designs, they mark up their sketches with more comments, similarly to how proofreaders mark up reports. Figure 1 shows an example of a sketch that a designer has annotated.

Annotations do not need to be only in written form. Other media such as audio and video can also be used. These media can be useful for annotations, because they can potentially carry more information than the written word alone. Besides words, speech also conveys the tone and mood of speakers and how fast they are speaking. Video adds body and facial gestures and eye gaze. Daft and Lengel characterize the capabilities of different media to carry different types of information the *richness* of the media [3].

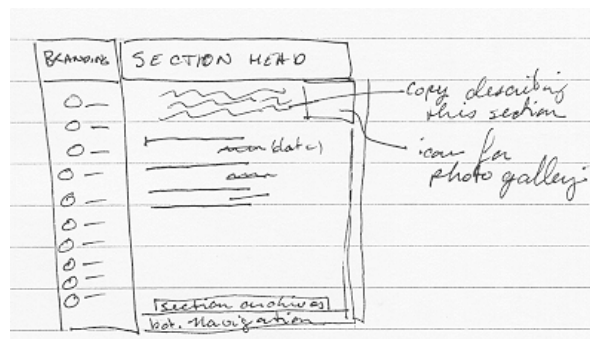


Figure 1. A sketch that the designer has annotated with some comments.

To determine how the medium of annotations affect their content and tone, I have implemented both voice and ink annotations in DENIM. I have also performed an informal evaluation with four participants, which showed a noticeable difference in the tone and style between ink and voice annotations.

RELATED WORK

There have been several studies comparing different media for annotating or authoring documents. Neuwirth et al [6] found that people who commented on a paper using voice annotations tended to focus on the purpose and audience of the paper, while those who typed their comments focused on substance. The reviewers of the paper also tended to use a more positive tone with voice annotations than with typed. Kraut et al [4] compared the use of face-to-face communication and e-mail for collaborative writing a document, and found that face-to-face is generally used more as the equivocality of the writing task increases, although the particular task and technology significantly affects the results. Chalfonte et al [2] performed a study where 12 MBA students annotated six brief texts using speech, audio, or both. They found that people tend to use speech for higher level issues of a document and text for lower level issues such as spelling.

IMPLEMENTATION AND USER INTERFACE

The DENIM interface has three main sections (Figure 2). The center area is an infinite canvas where the designer can sketch a web site. To the left is a slider that controls the zoom level of the canvas. At the bottom is a toolbox for holding tools. There are several types of tools: a hand tool for panning, a pencil and eraser for sketching and erasing the design, and a rubber stamp for creating reusable components. The behavior of the tools is similar to that of local tools [1]: users can “pick up” a tool by tapping¹ on it or hovering over it and pressing the space bar, and they can “drop” a tool by tapping on another tool

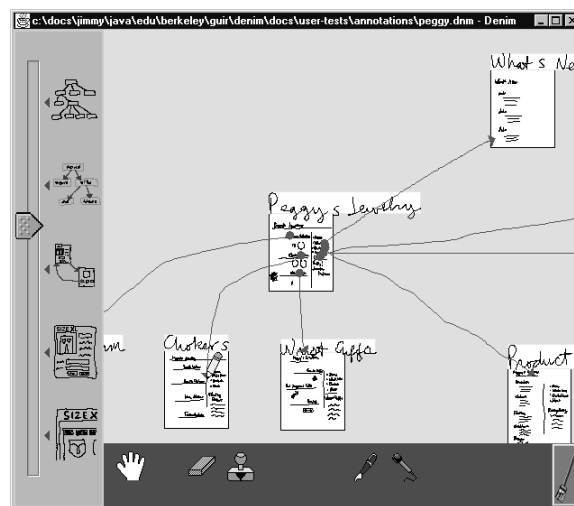
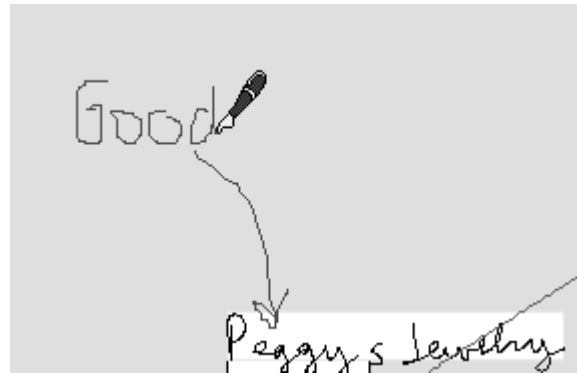


Figure 2. The DENIM user interface. The web site design inside DENIM was used in the user study.

¹ DENIM is focused on pen-based interaction rather than mouse-based, so I will use terminology such as “tapping” instead of “clicking.”

or by pressing the space bar.



I have added two more tools to the toolbox: a red pen for ink annotations, and a microphone for voice annotations.

Ink Annotations

Using the ink annotation pen is straightforward. The user simply taps on the red pen to pick it up, and sketches directly on the design. See Figure 3. Currently, all ink annotations are red so they are visually distinct from other parts of the design.

The tools for ink annotations and for normal sketching are separate because we want to make annotations visually distinct from the design, especially if the user is annotating someone else's design, and because DENIM recognizes and interprets certain marks while the user is sketching, such as creating a new web page, and we do not want this interpretation to occur during annotations.

Voice Annotations

To create a voice annotation, the reviewer first taps on the microphone to pick it up (Figure 4a). When he sees a part of the design that he wants to comment on, he taps it with the microphone. A separate “sound recorder” window pops up, allowing the user to record and playback his annotation (Figure 4b). When the user closes the window, a speaker icon appears where the user tapped (Figure 4c). Tapping the icon brings up the sound recorder window again.

EVALUATION

I performed an informal evaluation of the annotation features to determine whether a

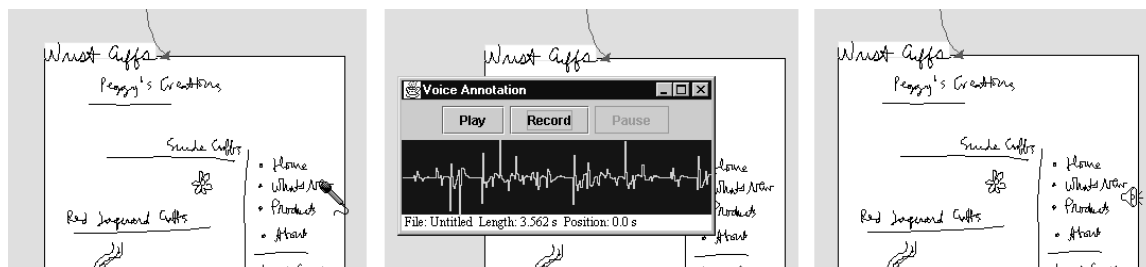


Figure 4. a) The voice annotation tool. b) Recording an annotation. c) The icon representing a voice annotation.

reviewer of a web site design would use ink and voice annotations differently and how. I

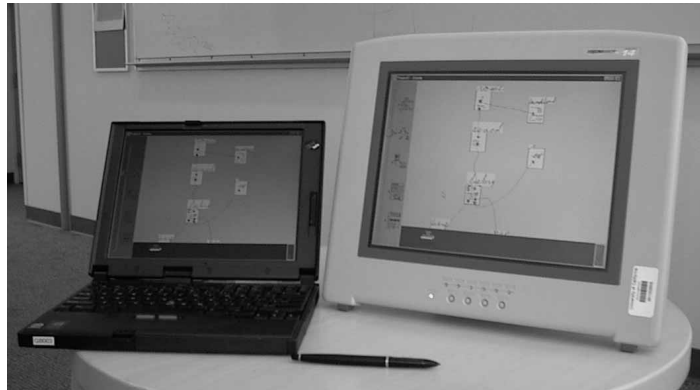


Figure 5. The laptop and the display tablet used for the informal evaluation.

had four participants for the evaluation, one professional web site designer and three UC Berkeley graduate students in computer science².

The hardware for the experiment consisted of an IBM ThinkPad 560Z 300 MHz Pentium II laptop running Windows NT 4.0, and an ITI VisionMaker Sketch 14 display tablet (Figure 5). The participants mostly interacted with the display tablet, although they could use the laptop's keyboard for shortcuts, and for the voice annotations they spoke into the laptop's built in microphone (we could not get an external microphone to work).

Participants were first asked to sketch a simple drawing in Microsoft Paint to get them familiar with using the display tablet. After I briefly demonstrated DENIM, I loaded an existing web site design into DENIM and asked participants to critically evaluate the design on such aspects as appropriate content, layout, consistent navigation, and consistent labeling, and to annotate the design with their comments with either voice or ink annotations. They then critically evaluated another web site design, using the medium that they had not used in the first evaluation. Each evaluation normally took 15 to 20 minutes, although one participant took as long as 30 minutes.

While the participants performed these tasks, I noted at which zoom levels they made their annotations, how many annotations were made per web page, and the tone and length of the annotations. I also recorded any usability comments that they had.

After all of the tasks, the participants filled out a questionnaire asking them which medium they preferred in which situations, how usable and useful the annotations were, and their overall satisfaction with the annotation features. The questionnaire also covered background information, such as basic demographics, their primary job responsibilities, and how much web site design experience they had.

Observations

The vast majority of ink annotations were made at either the storyboard or sketch levels, where the contents of each web page can be seen clearly. Participants used the site map level only to locate another web page to annotate. There were usually several ink

² Unfortunately, due to circumstances such as an overturned truck carrying raw sewage in San Francisco blocking traffic for six hours, I was unable to get more professional web site designers to participate in the study.

annotations for each element of a web page. A typical annotation was a complete sentence with an arrow pointing to a specific area of the page. Comments were usually a few words or a complete sentence. Similarly to inking a storyboard or sketch, participants could create a voice annotation for each element of a web page. The page, create a voice annotation. Occasionally, if there was something to add,

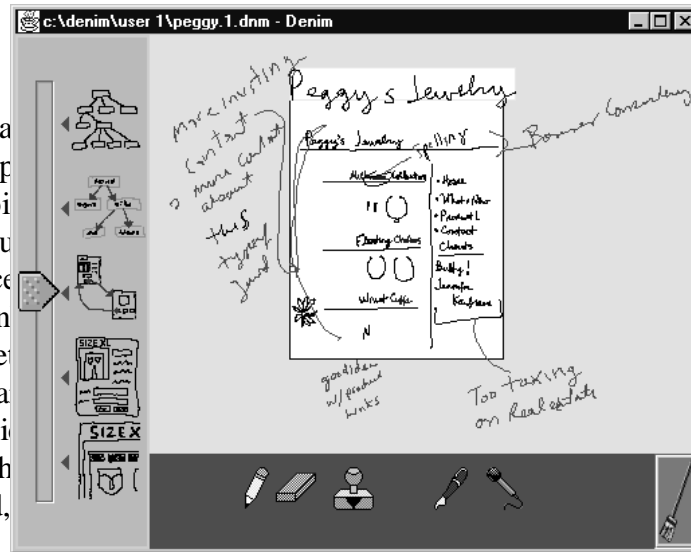


Figure 6. Typical ink annotations.

They almost always used complete sentences for each annotation. The main point was sometimes accompanied by additional explanation, which similar ink annotations lacked. For example, the web site designer had a couple of ink annotations that said, “Are these dynamic?” with arrows. When he made essentially the same comment with voice, he said, “This page should be dynamically built if there are a large number of listings. Also it makes it easier for updates.” This type of detail was not seen in ink annotations.

The tone between voice and ink comments was often different. One participant tended to preface any critical comment with a positive one within a voice annotation. Another participant ended many of his voice comments with “Thanks,” but not his ink annotations.

Feedback

The participants liked how the ink annotations were very easy and fast to create, but they complained about the amount of space that their annotations took up. Being able to hide or expand and collapse annotations could help with this problem. They also noted how hard it was to write legibly, even with the display tablet. Possible solutions to this problem include performing handwriting recognition and then cleaning up the ink, and allowing users to type in annotations. A couple of participants also said that they would like the ability to type in annotations in any case, even if writing on the display tablet was as easy as writing on paper.

Regarding voice annotations, the participants liked being able to make a lot of comments and not worry about the amount of space that they took. They also felt that some types of comments, such as overall conceptual issues, were easier to say than to write. However, there were also some drawbacks to voice annotations. Two participants felt that it was hard to structure their thoughts using only speech. One of them felt that written annotations allowed ideas to mature. He could write down his ideas, then later go back and easily review and edit them. Another participant said that with the present interface, there was no way to tell if one annotation was more important than another without listening to each one. He would like to use ink annotations to indicate how his audio annotations were related to each other. In effect, he would annotate his voice annotations with ink.

Overall, there was little consensus on which medium each participant preferred. Two people essentially had no preference, while one heavily preferred voice and one heavily preferred ink.

FUTURE WORK

There is still much work that can be done in this area. A more formal study comparing ink and voice annotations can be done, with many more participants in the professional web design field. The study can be expanded to include handwritten, spoken, and typed annotations. One can view handwritten annotations as being slightly richer than typed annotations, because some of the reviewer's personality and mood can come through in his handwriting. Whether this is an advantage that overcomes the problem of reading messy handwriting remains to be seen. Also, the study in this paper focused on how the reviewers used annotations. How the designers use the reviewers' annotations and how they react to the different annotation media is still yet to be studied.

The user interface of the annotations can also be improved. The study indicated that it is important to address the space issues of ink annotations. As I said in the previous section, being able to filter annotations, hide them, or collapse and expand individual annotations are possible approaches to this problem. The interface for the voice annotations can also be improved. Currently, users can only play and record annotations. They cannot rewind or fast forward to a particular spot, or append to an existing annotation. Users also cannot move a voice annotation icon once it has been created. All of these should be addressed.

There are also ways to more aggressively combine different media of annotations together. For example, voice annotations are currently tied closely to one page or one element of a page, simply due to where its icon is located in the design. DENIM could support the creation of "animated" annotations, where the reviewer speaks into a microphone while drawing on the design. This allows voice annotations to "address" more than one localized part of a design, while preserving the advantages of speech. Another idea is a little notepad associated with each voice annotation, so that someone listening to an annotation can write down notes about the annotation.

CONCLUSION

Annotations are clearly an important part of the design process. It is crucial for DENIM to support this capability. The study described in this paper shows that it is important to support multiple media for annotations. The different richness of ink and voice leads to the distinct roles that they play in annotations. There is still much more study to be done about how annotations can be best designed to fit how we use different media.

REFERENCES

1. Bederson, B. B., Hollan, J. D., Druin, A., Stewart, J., Rogers, D., & Profit, D. "Local Tools: An Alternative to Tool Palettes." In *Proceedings of User Interface and Software Technology: UIST 96*, pp. 169–170.
2. Chalfonte, B., Fish, R. S., and Kraut, R. E., "Expressive Richness: A Comparison of Speech and Text as Media for Revision." In *Human Factors in Computing Systems: CHI '92 Conference Proceedings*, pp. 21–26.

3. Daft, R. and Lengel, R., "Information Richness: A New Approach to Managerial Behavior and Organization Design." In *Research in Organizational Behavior*, vol. 6, 1984.
4. Kraut, R. E., Galegher, J., Fish, R. S., and Chalfonte, B., "Task Requirements and Media Choice in Collaborative Writing." *Human-Computer Interaction*, vol. 7, pp. 375–407.
5. Lin, J., Newman, M., Hong, J. I., and Landay, J. A., "DENIM: Finding a Tighter Fit Between Tools and Practice for Web Site Design." To be published in *Conference Proceedings on Human Factors in Computing Systems: CHI 2000*.
6. Neuwirth, C. M., Chandhok, R., Charney, D., Wojahn, P., and Kim, L., "Distributed Collaborative Writing: A Comparison of Spoken and Written Modalities for Reviewing and Revising Documents." In *Conference Proceedings on Human Factors in Computing Systems: CHI 94*, pp. 51–57.