Augmenting Real-World Objects: A Paper-Based Audio Notebook

Lisa J. Stifelman
Speech Research Group, MIT Media Laboratory
20 Ames Street, Cambridge, MA 02139
617-253-8026, lisa@media.mit.edu

ABSTRACT

The Audio Notebook allows a user to capture and access an audio recording of a lecture or meeting in conjunction with notes written on paper. The audio recording is synchronized with the user's handwritten notes and page turns. As a user flips through physical pages of notes, the audio scans to the start of each page. Audio is also accessed by pointing with a pen to a location in the notes or using an audio scrollbar. A small observational study of users in real settings was performed. The prototype did not interfere with the user's normal interactions yet gave reassurance that key ideas could be accessed later. In future work, automatic segmentation of the recorded speech using acoustic cues will be combined with user activity to structure the audio.

KEYWORDS: Speech interfaces, speech as data.

INTRODUCTION

While attending a lecture or meeting, a listener cannot write down a complete transcript of what is said. Oftentimes users may be missing critical information from their notes, want more detail for a particular topic, or need the ability to review the original material. While a cassette recorder can be used to capture a verbatim record, it is difficult to find the desired portions of speech, skim the recording, or correlate it with one's written notes.

The Audio Notebook *augments* a paper notebook, synchronizing the user's handwritten notes with an audio recording of the material. The user's natural activity (i.e., writing and page turns) serves as an implicit index into the audio. *Time* is mapped to *space*—the spatial layout of writing in the physical notebook enables rapid navigation in a time-dependent medium. Familiar objects like paper and pen are used for interacting with the audio rather than artifacts left over from analog devices, such as fast forward and rewind controls. The ultimate goal is to allow the user to *skim* both the written notes and audio recording.

APPROACH: WHY PAPER?

Previous systems for accessing time-based media from notes used computer-based input rather than actual paper and pen. NoTime linked notes written with a stylus on a notepad computer to an analog video recording [1]; Filochat linked writing on an LCD tablet to an audio recording [4]. The KeyRecorder[1] indexed an audio recording using typewritten notes on a laptop computer.

A goal of the Audio Notebook is to *augment* rather than *replace* the user's current notetaking tools. Paper and pen provide a portable, tangible, and ubiquitous way of capturing information [2]. The primary notetaking strategy for students and professionals attending lectures and meetings remains paper and pen. Paper documents have many advantages over digital ones—a sheet of paper can be quickly torn from a notebook, stuffed in one's pocket for easy access, or handed to a friend. Users often remember a mapping of *physical* locations in their notes to the desired information.

AUDIO NOTEBOOK PROTOTYPE

As the user writes in an ordinary paper notebook, the audio of a lecture or meeting is recorded digitally. Button controls for starting and stopping recording are activated by dipping the pen inside them (figure 1). After recording, audio can be accessed by space or by time. Playback can be started by pointing to a location in the notebook with the pen. Dragging the pen along an audio scrollbar navigates a time-line of the audio associated with each page.

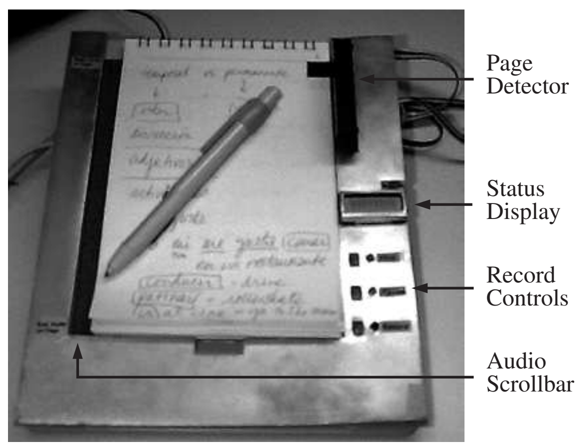

Fig. 1. Audio Notebook prototype.

The paper notebook is placed on top of a digitizing tablet; a U-shaped cover creates a slot for the notebook over the tablet's active area. The user takes notes using a cordless digitizing pen with an ink cartridge. The prototype can sense through a notebook of about 60 pages. Playback is triggered by the location and pressure of the pen. Playback begins a few seconds prior to the time offset of the user's selection. The current page number and state of interaction (e.g., playback time) are shown on a small LCD display.

[1] I developed this prototype in June of 1993.

LEDs are used to indicate when the device is recording or paused. The prototype is connected to a Macintosh Duo which handles recording, playback, and storage of pen data.

Page Detection

The prototype automatically detects the current page of the notebook during recording and playback. Bar codes were considered, but the readers are large and require manual input by pointing or swiping. Instead, a specialized page detector was developed (figure 2). Page numbers are printed in a six bit binary code along the side of each page using black and white squares. The code is read using six optical sensors angled down at the page.

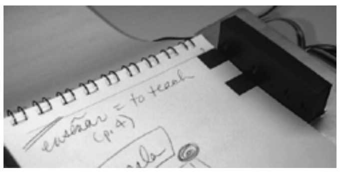

Fig. 2. Audio Notebook page detector.

USER STUDY

A small study was performed to observe use of the audio notebook prototype in real settings. Users were selected who had a need to take notes and were strongly motivated to review them. The study had two parts—profiles of user's current notetaking habits, and use of the audio notebook for taking and reviewing notes.

User Notetaking Profiles

To learn about the notetaking habits of different users and determine if their notetaking style changed when using the Audio Notebook, the following types of users were interviewed: a language student, a computer science student, a user studies professional, and a lawyer. The lawyer and technical students needed very detailed notes. The lawyer often attempted to capture a verbatim account; obtaining missing information from a client is usually not possible or advisable. The user studies professional recorded all his interviews on cassette tape and manually transcribed them. He took few notes, jotting down his own thoughts, rather than what subjects said. The language student also took few notes; the classroom sessions involved mostly speaking and interacting with others.

Taking Notes with the Audio Notebook

Three of the users interviewed participated in the second part of the study—use of the audio notebook in real settings. Two subjects used the prototype during one of their class sessions, and one during a group meeting.

Prior to the notetaking session, each user was only told "the device synchronizes your notes with an audio recording." Given only these brief instructions and no prior experience with the device, it was surprising that users' notetaking was affected. The language student wrote larger and took more notes, "almost unconsciously" annotating points for later access. The user studies professional tried to anticipate whenever something important was said by constantly pausing and restarting recording. Part way through the meeting, he changed his strategy, keeping recording on at all times and using the notes to mark indices. A social implication of his original strategy was that others noticed his actions and said "why are you pausing me!"

Reviewing Notes with the Audio Notebook

Playback by selection with the pen was intuitive. However, users did not like the way playback started automatically whenever they turned a page. Users, at least inexperienced ones, tend to want tight control over starting and stopping audio playback [3].

The purpose of the audio scrollbar was not immediately apparent. The scrollbar is intended for fine-grained control in cases when notes are sparse. A slot in the middle of the scrollbar provided an obvious place to slide the pen, but users thought its function was redundant with the ability to play directly from the page. However, after using the scrollbar for a short period, it was noted as a good way to quickly "run down a page." Displaying a correlation between time and space, such as lighting up a point in the scrollbar whenever the user selects on the page, could enable the user to adjust the starting point of playback.

Another means of playback was suggested by the language student. She wanted to distill a "key note" audio summary from writing that was circled or underlined in her notebook.

CONCLUSIONS AND FUTURE WORK

When asked what she liked the best about the audio notebook one user said "it doesn't alter the fundamental way I take notes, or my interactions in class" yet she found it reassuring that key ideas would be accessible later on. This goes back to the assumptions about using paper and pen and why this is so important. Rather than simply replacing real world artifacts, we can augment them, providing computational capabilities not previously possible. Perhaps the Audio Notebook is more than just an augmentation of existing paper media, but the creation of a new artifact through the seemless integration of paper and audio.

This paper has presented an approach for "user structured" audio. Future work will attempt to combine user activity with acoustic cues (e.g., changes in pitch and pauses) for automatically segmenting a speech recording.

ACKNOWLEDGMENTS

Thanks to Wayne Burdick, Jacob Tuft, and Tai Mai for hardware development. Thanks also to Chris Schmandt, Barry Arons, David Reed, Debby Hindus, Andrew Singer, Hiroshi Ishii, and everyone at Interval Research Corp.

REFERENCES

1. Lamming, M.G. Towards a Human Memory Prosthesis. Tech Report #EPC-91-116. Rank Xerox EuroPARC, 1991.

2. Newman, W. and Wellner, P. A Desk Supporting Computer-Based Interaction with Paper Documents. *In Proc. CHI '92*, pp. 587–592. ACM, 1992.

3. Stifelman, L.J., Arons, B., Schmandt, C. and Hulteen, E.A. VoiceNotes: A Speech Interface for a Hand-Held Voice Notetaker. *In INTERCHI '93*, pp. 179-186. ACM, 1993.

4. Whittaker, S., Hyland, P. and Wiley, M. Filochat: Handwritten Notes Provide Access to Recorded Conversations. *In CHI '94*, pp. 271–277. ACM, 1994.