NoteLook: Taking Notes in Meetings with Digital Video and Ink

Patrick Chiu[1], Ashutosh Kapuskar[1], Sarah Reitmeier[2], Lynn Wilcox[1]

[1]FX Palo Alto Laboratory, 3400 Hillview Ave, Bldg 4, Palo Alto CA 94304, *lastname*@pal.xerox.com

[2]University of Michigan, School of Information, Ann Arbor, MI 48109, sreitmei@umich.edu

ABSTRACT

NoteLook is a client-server system designed and built to support multimedia note taking in meetings with digital video and ink. It is integrated into a conference room equipped with computer controllable video cameras, video conference camera, and a large display rear video projector. The NoteLook client application runs on wireless pen-based notebook computers. Video channels containing images of the room activity and presentation material are transmitted by the NoteLook servers to the clients, and the images can be interactively and automatically incorporated into the note pages. Users can select channels, snap in large background images and sequences of thumbnails, and write freeform ink notes. A smart video source management component enables the capture of high quality images of the presentation material from a variety of sources. For accessing and browsing the notes and recorded video, NoteLook generates Web pages with links from the images and ink strokes correlated to the video.

Keywords

note taking, meeting capture, video applications, multimedia applications, electronic meeting support, electronic notebook, pen computing

1. INTRODUCTION

Multimedia notes and records of meetings provide many benefits over their traditional paper counterparts. A video recording of a meeting allows people to review a meeting that they have attended or to catch up on a meeting that they have missed. Note takers do not have to worry about falling behind or missing details, and electronic notes can serve as indexes for searching and browsing the recorded video. With increasing use of visual material in presentations, providing the ability to capture the slide images unburdens note takers from the tedious task of copying down information from presentation slides.

While there exist a number of research systems that employ various combinations of pen-based devices in conjunction with audio and/or video ([1], [2], [8], [10], [11], [12], [14], [15], [16], [17], [18], [19]), the use of digital video as a medium that users in meetings can interact with and incorporate into their notes has not been explored.

We have designed and built a highly interactive client-server system called NoteLook to support multimedia note taking in meetings with digital video and ink. NoteLook is integrated into a conference room at our lab, which is equipped with computer controllable video cameras, video conference camera, and a large display rear video projector. The video cameras provide images of the speaker and room activity, and the rear video projector provides images for capturing the presentation material in a variety of forms. Images can be interactively and automatically incorporated into the notes taken with the NoteLook client application on wireless pen-based notebook computers. The video is transmitted over a wireless network and stored for later retrieval. At the end of a meeting, participants can take away paper printouts of the notes with images and handwritten ink. NoteLook also has a facility to publish notes on the Web with links to the recorded video.

To evaluate and refine the design of NoteLook, we tested a prototype system with subjects taking notes in several types of real meetings: presentations, staff meetings, and Japanese classes. The system provided seamless capture of presentation material in a variety of forms and was unobtrusive to the speakers and other participants in the room. We found the application to be flexible enough to support different styles of note taking. Findings from the user tests led to refinements in the system architecture and a new automatic note-taking feature.

This paper is organized as follows. We begin by pointing out challenges and tradeoffs in section 2. Next, we describe the NoteLook application in section 3. Then we go into the system architecture in section 4. We discuss the design history including user tests and refinements in section 5. We outline related work in section 6. Finally, we conclude with section 7.

2. CHALLENGES AND TRADEOFFS

There are several important challenges and tradeoffs in working with digital video and ink for note taking and meeting capture. We point them out and explain our approaches for dealing with them.

2.1 Lightweight/Heavyweight

For note taking and meeting capture, it is desirable to have a lightweight system that does not hinder the note taker from paying attention to the meeting and gets in the way of the speaker and other participants. Digital video is a heavyweight medium because a substantial infrastructure is required to obtain adequate quality images of the room activity and presentation material. It is necessary to separate the system into components in order to provide the note taker with a lightweight application. In contrast, when working with digital audio and ink, a stand-alone system is sufficient (e.g. Audio Notebook [14] or Dynomite [18]). To achieve a lightweight form factor for the note-taking application component, wireless pen-based notebook computers can be used, and the heavyweight components that process the video and manage the video sources can be hidden away from the users by placing them outside the meeting room.

2.2 Versatility/Fidelity

Video is a versatile medium because it has the ability to capture the room activity as well as the presentation material in a variety of forms. The video feed from a rear video projector can be used to capture images of the presentation material regardless of the type or format used by the speaker. However, a noticeable loss in fidelity may occur due to signal noise. For example, by the time an image of a PowerPoint slide travels from a PC's video output through the plumbing (which may contain various splitters and scan converters) and reaches the rear projector, the captured image is degraded to a level that sometimes makes it difficult to read the text on the slide. A method to address this *versatility/fidelity tradeoff* is to use a smart video source management component to find the highest fidelity source available for capturing images rather than just always taking the source from the rear video projector.

2.3 Interactive/Automatic

Note taking by nature is a highly interactive activity. For meeting capture, electronic notes with time-stamps from a scribe or meeting participants provide a good way to index the recording of a meeting. We found from user tests that when people taking notes are given the capability to incorporate images into the note pages, they sometimes want to take many images (e.g. collecting a complete set of the presentation slides and annotating them). To support this style of note taking better, it is helpful to have a feature that automatically captures slide images into the note pages as the speaker puts up new slides.

3. THE NOTELOOK APPLICATION

In this section, we describe how the NoteLook application is used to support note taking and meeting capture and how it addresses the challenges and tradeoffs outlined above. The NoteLook application operates in a media enriched conference room equipped with video cameras, a video conference camera, and a large display rear video projector. Users take notes on wireless pen-based notebook computers which may be freely positioned around the room. They can interactively or automatically snap images from the cameras or rear projector into their note pages. For later access, NoteLook generates Web pages of the notes with links to the recorded video.

3.1 Media Enriched Conference Room

In our laboratory, we have a conference room equipped with three room cameras and a large 120-inch screen rear video projector. Figure 1 shows the conference room. The podium has controls that allow the speaker to select a source for the rear projector. Video images of the presentation material are fed into the rear projector from a PC workstation, a user-supplied laptop, a document camera, a VCR, or a DVD player. A room camera is used to capture images of content on the whiteboard. Audio is captured by six ceiling microphones, combined into a single audio stream and stored together with the video streams. These streams are transmitted over the wireless network to the note-taking devices. Notes are taken on wireless pen-based notebook computers (Fujitsu Stylistic 1200 with an 8-inch active matrix VGA display). An ink jet printer is available to produce hardcopies of notes.

3.2 Interactive Note Taking

Meeting participants use the pen-based computers in the conference room to take notes. See Figures 1 and 2. Since the computers have the small form factor of a notebook and are wireless, users may sit comfortably anywhere in the room without disturbing their neighbors. The note-taking software on pen-based computers is the NoteLook client application. It runs on Windows 95 and is a descendant of the Dynomite [18] ink and audio note-taking application.

Figure 1. Conference room with NoteLook.

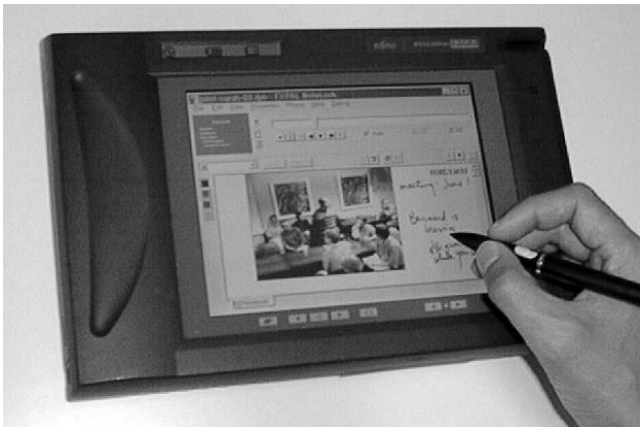

Figure 2. NoteLook client running on a pen computer.

The NoteLook client has a main note page area for taking notes (see Figure 3). The user can write freeform notes and snap in images from the video. During note taking, the user can view live video in the small window on the upper left corner. This video window also displays the recorded video during playback. Next to the video window are three buttons for interacting with the video. The top button changes the video channels. We normally run the system with two channels: one for the room activity from a pre-selected room camera and one for presentation material shown on the rear projector. Usually, the pre-selected camera is pointed at the speaker at the podium. Sometimes an additional channel showing a wide angle shot from the back of the room is used. The middle button snaps a thumbnail of the image from the video window into the margin of the note page. When a sequence of thumbnails is snapped, they are placed one below another. The bottom button snaps in a large image as the note page background. A newly snapped background image overwrites an existing background image on the page.

There is also a set of standard VCR-type controls. Pressing the RECORD button makes a connection to the NoteLook server and initiates video recording and transmission to the client. Above the VCR controls is a timeline with a pointer for indicating the current time position of the video.

At the top right corner of the note page are buttons for previous page, next page, and new page. On the left is a palette of four pen colors for writing notes and annotations. Underneath the video window is a list box for entering keywords, and further to the right is a set of four buttons for query and retrieval. Details of the query and retrieval features are similar to those of its predecessor Dynomite [18], which we will not discuss further in this paper.

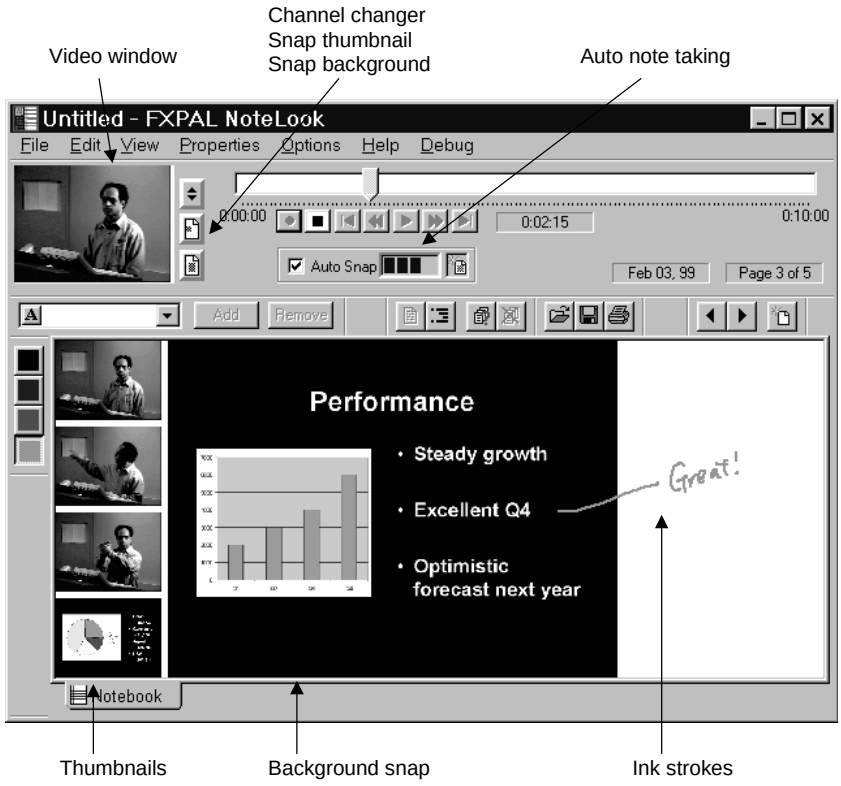

Figure 3. Screen shot of the NoteLook client application.

3.3 Automatic Note Taking

NoteLook has a facility for automatic note taking. In this mode, when the presenter puts up a new slide on the rear projector, it is automatically detected and snapped in as a background of a new page, and this page is appended to the stack of note pages. Also, a sequence of thumbnails from the room cameras is placed in the margin of that page. When the user turns to that page, she can annotate it with ink. This feature is helpful for presentations with many slides. In our experience, it is common to see 20 slides in a presentation and we occasionally have talks with over 50 slides.

To use this feature, the user checks the checkbox for Auto Snap (See Figure 4). As a new slide comes in, a progress indicator with animation gives user feedback. An icon is displayed to indicate that there are new pages. Pages are not turned automatically because the user may still be working on the current page.

For slides this works well. The reason is that slide changes can be detected fairly accurately by simply looking at image differences. For more dynamic types of presentation material such as video clips, the user is better off snapping in the images manually.

The sequence of thumbnails consists of images from the room cameras taken at evenly spaced intervals during the time that the corresponding background slide is on the screen. Video and image analysis techniques may also be employed to find interesting images to incorporate into the notes.

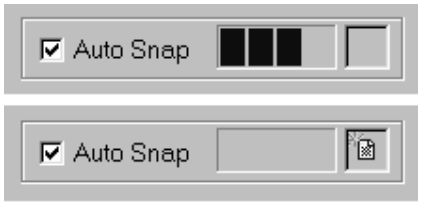

Figure 4. Automatic note taking indicator. The top picture shows a new slide coming in. The bottom picture shows that there are new pages.

3.4 Accessing Notes and Recorded Video

NoteLook notes and their associated video recordings may be accessed and played back on the pen computer. The video plays back in the video window (in the upper left corner of Figure 3). The thumbnails, backgrounds, and ink strokes are all time-stamped during note taking and can be "played" by selecting them and pressing the PLAY button on the VCR controls. For example, playing a thumbnail would play the recorded video at the time that the thumbnail was snapped in by the note taker.

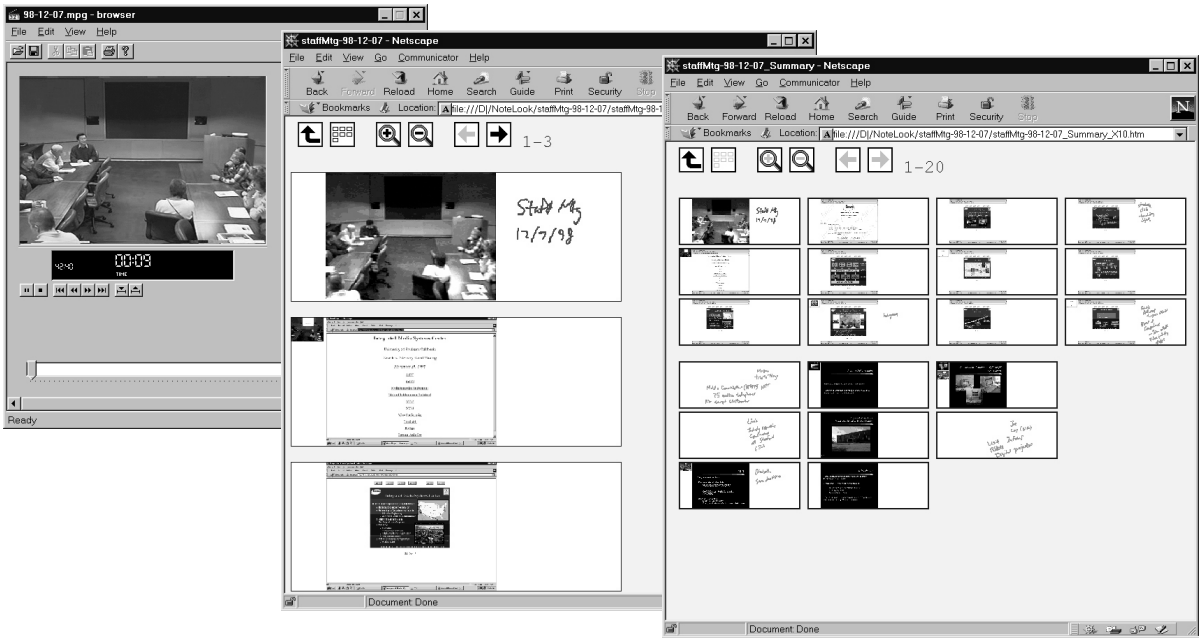

Figure 5. Web access to NoteLook notes and recorded video. In the center is a Web page showing three NoteLook note pages, on the right is a table of contents page, and on the left is the Metadata Media Player.

To promote wider access to the notes and recorded video, NoteLook creates Web pages with links from the images and ink strokes correlated to the video. A sample is shown in Figure 5. A menu command in the application generates the HTML pages. Clicking on the images and ink strokes plays the video in a separate application window. NoteLook Web pages are integrated with a video playback application developed at our lab called the Metadata Media Player [5].

Additionally, there are several standard navigation features. These are straightforward and we will only give a brief description. A top level page lists available NoteLook notes. For each session, a single table of contents page shows reduced images of all the pages. Clicking on a reduced image of a page brings up that page. Each web page may be zoomed in or out with a range of five different magnification levels.

4. SYSTEM ARCHITECTURE

The NoteLook architecture is designed to address the challenges and tradeoffs described in section 2 above. NoteLook provides a lightweight client application for the note takers and performs most of the processing on servers outside the conference room. To deal with the versatility/fidelity tradeoff, NoteLook has a smart video source management component that identifies and directs the highest fidelity source to the server for capturing. While being a highly interactive application, NoteLook also supports automatic note taking.

The NoteLook architecture is auto-configurable, extensible, and scalable. Video channels may be added or removed without having to modify existing NoteLook clients and servers. This is accomplished by using resource discovery techniques.

The various components of NoteLook are shown in Figure 6. The key pieces are the NoteLook clients, servers, and switchers for video source management. We describe the interplay of these along with other components in more detail below.

4.1 NoteLook Client

The features of the NoteLook client application have been described earlier in section 3. We now explain how it works with the other parts of the NoteLook system. The client devices communicate with the servers over a wireless network with a base station located outside the conference room (see Figure 6). The base station and servers are connected to a LAN.

Each server corresponds to a video channel. To receive video from the servers, a client locates and connects to the servers by means of a resource discovery mechanism. The connection process is initiated when the user presses the RECORD button on the client. The client then broadcasts a request for service, the servers respond and identify themselves, and connections are established. The clients do not save the video locally on the pen computers with its note pages because of the large amount of disk space required by video. Instead, the servers save the video to a database on the network.

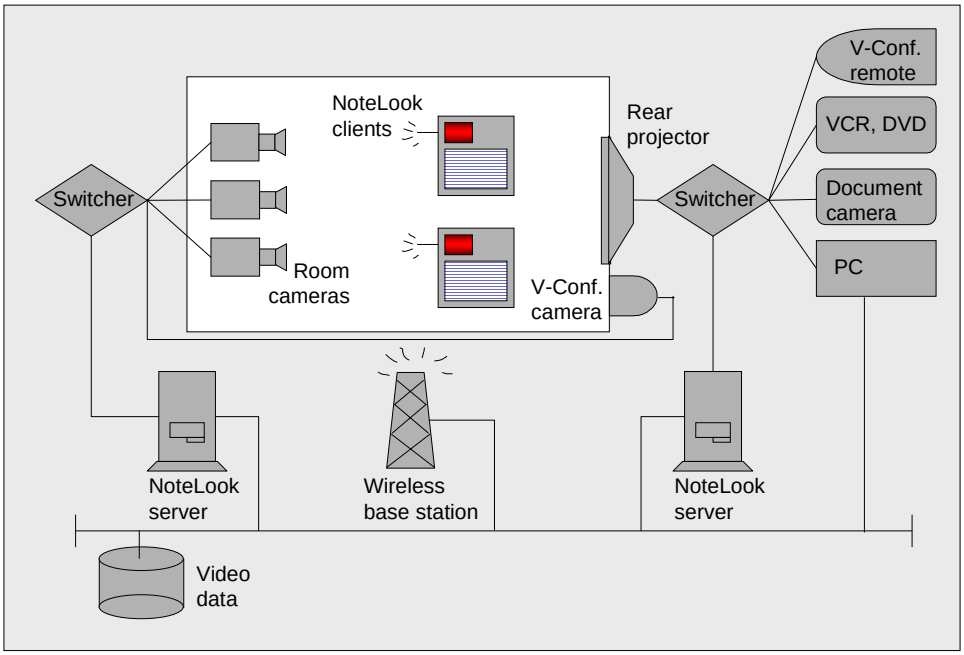

Figure 6. NoteLook architecture. The white area shows the components inside the conference room.

With the resource discovery mechanism, channels can be easily added, removed, and reconfigured, even while the clients are running. Sometimes it is desirable to reconfigure the channels for different situations, and to do this within a minute or two from the start of a meeting. For example, a presentation may be followed by a discussion. During the presentation, two channels can be used: one with a view of the podium and the other the rear video projector. In between sessions, a third channel can be added, and these can be reconfigured to use the three room cameras to cover people talking from various parts of the room during the discussion.

4.2 NoteLook Server

The NoteLook servers take video and audio inputs, process them, transmit the output to the NoteLook clients, and store the multimedia data for later retrieval. One server is used per channel. The servers run on Windows NT PCs with video capture cards and sound cards installed. Digital video is captured at 15 or 30 frames a second with a resolution of 640x480. Video data is stored on the network for playback.

Using multiple NoteLook clients at the same time is supported. The servers record and transmit only when there is at least one client running in the RECORD mode. Thus, a meeting may produce different sets of note pages associated with the same video recording.

To conserve bandwidth on the wireless network, the server transmits video at a highly reduced frame rate. Since the user is present at the meeting and paying attention to the live action, it is not necessary to have full motion video on the NoteLook client. The frames have to be updated frequently enough to capture slide changes and pictures of activity in the room. Our experience is that one frame per 2 seconds is adequate for note taking.

4.3 Smart Video Source Management

While each channel corresponds to a server, it does not correspond to a single video source but rather to a set of video sources managed by switchers. In the current setup, there are two channels: one for the presentation material and another for the room activity. The presentation material channel corresponds to the set of sources:

{*Rear Projector, Document Camera, VCR, DVD, PC, (User-Supplied Laptop), (Video Conference Remote)*}

The room activity channel corresponds to the set:

{*Camera 1, Camera 2, Camera 3, Video Conference Camera*}

Different channels may be mapped to different sets of sources. For each set of sources, a switcher directs the appropriate source to the server for its channel. The video data flow model is shown in Figure 7.

This design gracefully deals with the versatility/fidelity tradeoff. On the presentation material channel, the switcher identifies the actual source that is being shown on the rear projector and selects the highest fidelity source available for capturing images. For example, when slides are shown on the rear projector from PowerPoint running on the PC, the server will get the high-fidelity images from the PC display by capturing a screen bitmap (not the PC's video

monitor signal and not the rear projector video signal). In the case when a speaker supplies her own laptop, the server gets its images further downstream from the video signal of the rear projector with some unavoidable degradation in fidelity. The switchers operates in real time. The result is that the best obtainable images are always captured while the smart video source management is hidden from the user.

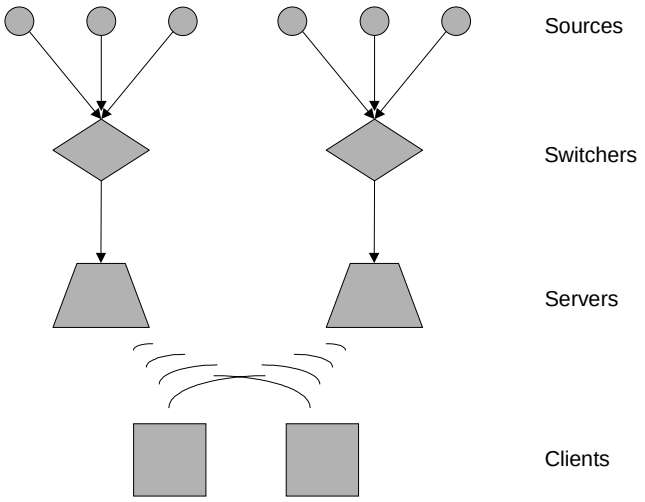

Figure 7. Video data flow.

For the room activity channel, the cameras in the conference room are computer-controllable Sony video cameras. A camera is chosen and positioned before a meeting, and sometimes there is a person in the backroom controlling the cameras. A switcher identifies the preset camera and directs the composite video signal to be digitized by the NoteLook server equipped with a video capture card. When there is a single user (*i.e.* a meeting scribe), it is possible to modify the system so that the user may select and control the room cameras while taking notes during a meeting. In general when there may be multiple note takers, the cameras should not be controllable by individual note takers to avoid conflicts.

4.4 Automatic Note Taking Component

The NoteLook automatic note taking component, AutoNote, is a separate piece of software that runs on one of the servers. It determines a set of images consisting of one background image and a sequence of *n* thumbnail images that will be used by the client to create a new page. Our default channel for the background image is the presentation material channel, and for the thumbnails the room activity channel. The default for *n* is 4, which fills the height of a note page. AutoNote monitors a buffer kept on the presentation channel server to learn about the presence of new slides. It also monitors a buffer for the room camera and chooses the sequence of thumbnail images.

We now explain how the images are chosen for an automatically created page. The NoteLook client periodically checks with AutoNote to see whether a new page is ready to be created. From our experience, we found that 6 seconds is an acceptable granularity for checking for fresh slides. While the NoteLook client in the Auto Snap mode (section 3.3) indicates to the user when a new page with a fresh slide has "come in," it does not actually create the page until the user goes to that page. This delay is what makes the selection of several thumbnails from a single room activity channel possible. If there were no delay, we would have to choose a room shot at the instant the fresh slide was put up. With the delay, we have a time span with a start time when the fresh slide was put up, and an end time when the slide was taken off. The heuristic for choosing the sequence of thumbnails in this time span is as follows: we take the video frames so that they are evenly spaced in the time span. Other ways of choosing images may be used, such as applying video and image analysis techniques for finding interesting images to incorporate into notes.

4.5 Web Access to Notes and Recorded Video

We describe how NoteLook generates the Web pages with links from the images and ink strokes correlated to the video. Each HTML page consists of three NoteLook page images, and each of these is overlaid with an HTML image map. This links the areas under the individual thumbnails, backgrounds, and ink strokes to the corresponding times in the recorded video. Clicking on these areas calls the Metadata Media Player [5] to start playing the video from the appropriate time. The recorded video is stored and accessed from a central video database. A separate page is also produced for each of the five magnification factors. Finally, a table of contents page containing small images of all the pages with links to their Web pages is produced.

5. DESIGN HISTORY

The current design of NoteLook came about through a process of rapid prototyping, user testing, and refinement.

5.1 Initial Design

The initial design of NoteLook was fairly basic without automatic note taking and smart video source management. We wanted to rapidly prototype a working system and have users test it before venturing into more sophisticated features. For the client, we re-used code from the Dynomite [18] audio and ink note-taking application by adding capabilities for handling video and interactively incorporating images into the note pages. Then we built the servers and wired them to the conference room video sources. We manually selected a room camera for the room activity channel and hooked up the rear projector video feed for the presentation material channel.

5.2 User Study

Next, we conducted a user study to evaluate this basic NoteLook system and to observe how people would take notes given the digital video and ink capabilities. We also wanted to observe the impact (or non-impact) of the system on the speaker and other participants in the room. Our plan was to study the regular use of NoteLook during real events with real content: presentations, staff meetings, and Japanese classes.

As a preliminary step, we ran 11 pilot tests with researchers from the NoteLook project to check for bugs in this complex client-server system. Following the pilot tests, the actual study took place during a six-week period, in which 4 subjects took notes in 13 sessions. A breakdown of the meetings is shown in Table 1.

Meeting Type	Sessions	Subject
Presentation	2	S1
	4	S3
Staff meeting	3	S2
Japanese class	4	S4

Table 1. Breakdown of user test sessions.

The subjects consisted of one administrative assistant, one system administrator, and two researchers. None were associated with the NoteLook project. Each subject received 20 minutes of training. A single user took notes in each session. The task given to the user was to take notes for himself and also for colleagues who may miss the meeting. With this task, the user was encouraged to take notes but was not assigned the formal role of a scribe. After each session, a printout of the notes was given to the user.

Observations and data were collected from each session by interviews, direct observation, user logs, NoteLook files and printouts. Before each session, the subject was asked if he had any questions about NoteLook. After each session, the subject was interviewed about his general impressions of the system, how they used the system, the interface, and the printouts of the notes. Also, presenters and people sitting next to the subject during the presentation were briefly interviewed.

One result we found is that NoteLook supports seamless capture of content from a variety of presentation material types and formats. Speakers used the following: PowerPoint slides, PowerPoint slides embedded with video, a custom JPEG viewer, Web pages, overhead slides on the document camera, whiteboard, and video from the VCR. Of the 6 presentations, 5 were external speakers and 4 brought their presentation material on laptops (2 MACs, 2 PCs). The laptop video output signal was fed into the rear projector via a connector on the podium. One speaker had PowerPoint slides on a floppy disk that were loaded into the conference room PC. Three speakers had overhead slides that were displayed on the document camera. In all cases, NoteLook enabled users to capture the visual material easily and seamlessly.

We also found that the system was minimally intrusive to the speaker and other participants. Out of the 7 speakers interviewed, only 2 even noticed NoteLook and none of them found it distracting. One of the presentations was a job interview talk, and another was by a former head of the laboratory. These events are of such serious nature that intrusions would not be tolerable.

With any kind of recording technology, there are privacy concerns. We have a policy at our lab of asking speakers for permission to be videotaped and none of them objected. Members of the laboratory are familiar with the conference room and know that the cameras are live.

Two distinct note-taking styles were observed. We call these two styles *image-centric* and *ink-centric*. The image-centric note takers (S1 and S3) took many background snaps and thumbnails, and produced a set of note pages that resembled a full set of the speaker's slides with annotations. In contrast, the ink-centric note takers (S2 and S4) had more handwritten ink supplemented by some images, and fewer pages of notes. It is possible that other styles may emerge as more users take notes with the system over long periods of use.

All the users said that they would use NoteLook again. They also raised a few issues: S3 expressed that having high quality images is important, and S4 wanted the resulting notes available in electronic form. There is also evidence that at least two of the subjects used their notes afterward. S4, who took notes in Japanese classes, referred back to his NoteLook printouts from the week before during one of the sessions, just as he would have referred back to notes taken in his paper notebook. S1 also transcribed his notes for research purposes, just as he would have normally, but found the visual images added more context.

5.3 Refinements to System Design

The user tests made it clear that there is a versatility/fidelity issue. The tradeoff is that while NoteLook can capture presentation materials in a variety of forms, image quality is not the highest if we only rely on the rear projector. This led us to develop the smart video source management component of the system.

The desire by test subjects to have electronic copies of notes prompted us to deliver easy Web access by having NoteLook generate HTML pages with links to the recorded video. Further user feedback suggested the zoom feature and table of contents page.

Automatic note taking was designed to help the image-centric note taker. In a presentation, it is common to see 20 or more slides and it requires much work by the user to manually snap in the images. The Auto Snap feature relieves the user from this repetitive task.

6. RELATED WORK

Related work may be broadly categorized into meeting capture systems, stand-alone ink and audio notebooks, video annotation systems, and PDAs with digital cameras.

6.1 Meeting Capture Systems

NoteLook represents an advance over earlier integrated multimedia note-taking systems for meeting capture. None of these systems allow users to incorporate interactively live video and images from cameras and presentation material into the notes.

We-Met [19] is a pen-based meeting support system with a shared drawing surface. Meeting events on the pen-based shared display can be recorded and played back. This work also reports on a study of retrieving specific items in videotapes of meetings.

The STREAMS [3] system uses video cameras in an auditorium for presentation capture. These cameras are also used to capture any presentation material that is shown. This method has problems when activity in the room obscures the display. Skewing is another problem with employing an externally mounted camera to capture presentation content. Note taking during a presentation is not supported, although the captured video streams can be annotated later during review by adding text comments.

The Forum [7] system uses video for distributed presentations. Everyone, including the speaker, sits in front of a workstation during a presentation. Slides have to be prepared in a specified format. The slides can be annotated with text and marks drawn with a mouse, but the video images cannot be annotated.

A confederation of multimedia tools for meeting capture and salvage [10] uses the Tivoli ([11], [12]) system running on a LiveBoard [13]. Slides created and annotated in Tivoli are captured together with the audio during a meeting. Tivoli is designed to support working meetings. The group notes to the meeting are the Tivoli whiteboard pages with ink strokes, graphical objects, and type text notes "beamed up" from a laptop. The audio is correlated to each time-stamped object on the Tivoli page.

In the Classroom 2000 ([1], [2]) system, the audio from a lecture is recorded and slides with annotations are made available to students after class. The presenter must make a special effort to prepare the slides in a standard graphics format. The slides are displayed on a LiveBoard and note taking is done with PDA devices pre-loaded with slides. These notes are later synchronized to the audio and the slides that have been annotated by the professor on the LiveBoard.

NotePals [4] is a lightweight note sharing system. Notes and taken in classrooms and at conferences on stand-alone pen-based PDAs (Palm Pilots) and then the notes are uploaded to a shared repository on the Web. There is no support for audio or video.

6.2 Stand-Alone Ink and Audio Notebooks

There are several stand-alone ink and audio note-taking systems. These systems have neither image capture nor integrated video capabilities. NoTime [8] pioneered the idea of time-stamping ink strokes and correlating them to recorded audio or video. Filochat [17] is a PC computer with a pen tablet in which audio is indexed to handwritten notes. FXPAL Dynomite [18] is a predecessor of the NoteLook client application. Audio Notebook [14] uses a special device for taking notes on paper, which are correlated to an audio recording.

6.3 Video Annotation Systems

Somewhat related are video annotation systems. These systems are limited by their design based on using videotapes rather than digital video. None of these systems allow interactive incorporation of video images into the notes. VideoNoter [15] displays and synchronizes different streams of activity (video, figures of whiteboard drawings, text), but requires post-production to transcribe text from the audio or extract drawings from a whiteboard. Vanna [6] and EVA [9] are text-based video annotation systems. Marquee [16] is a pen-based system for making ink annotations while watching a videotape. A later version of Marquee has modifications to take time-stamps on digital video streams from the WhereWereWe [10] multimedia system.

6.4 PDAs with Digital Cameras

There exist commercial products that allow digital photos to be annotated with digital ink. The small display size on these devices makes them unsuitable for taking extensive notes with pictures in meetings and presentations. Sharp Power Zaurus is a PDA with an attachable digital camera. Digital photos can be taken and linked to handwritten notes. Nikon Coolpix 300 is a digital camera with a stylus for making annotations on top of digital photos. These systems also allow audio snippets to be linked to image or note objects.

7. CONCLUSION

We have presented NoteLook, a highly interactive client-server system integrated into a media enriched conference room for note taking and meeting capture. We identified several challenges and tradeoffs and described our approaches to them. We also provided an account of the design history that involved rapid prototyping, user testing,

and refining the system. Our efforts produced several novel components and features: a lightweight wireless pen-based client that allows users to interactively incorporate images into their notes, a smart video source management component for obtaining the highest fidelity images, and an automatic note taking component. In the future, we plan to acquire more usage experience with NoteLook and integrate it into our meeting work practice.

8. ACKNOWLEDGMENTS

We thank John Boreczky and Andreas Girgensohn for help on integrating NoteLook Web pages with the Metadata Media Browser, Andreas for help on the screen capture component, John Doherty for video support, Sara Bly for help with the application design and user studies, our four user study subjects for their participation, and readers of this paper for their valuable comments.

9. REFERENCES

[1] Abowd, G. D., Atkeson, C. G., Brotherton, J., Enqvist, T., Gulley, P., and LeMon, J. Investigating the capture, integration and access problem of ubiquitous computing in an educational setting. *Proceedings of the CHI '98 Conference*. ACM Press, pp. 440-447.

[2] Abowd, G. D., Atkeson, C. G., Feinstein, A., Hmelo, C., Kooper, R., Long, S., Sawhney, N., and Tani, M. Teaching and learning as multimedia authoring: the classroom 2000 project. *Proceedings of the ACM Multimedia '96 Conference*. ACM Press, pp. 187-198.

[3] Cruz, G. and Hill, R. Capturing and playing multimedia events with STREAMS. *Proceedings of the ACM Multimedia '94 Conference*. ACM Press, pp. 193-200.

[4] Davis, R., Landay, J., Chen, V., Huang, J., Lee, R., Li, F., Lin, J., Morrey, C., Schleimer, B., Price, M., and Schilit, B. NotePals: Lightweight note sharing by the group, for the group. *Proceedings of the CHI '99 Conference*. ACM Press, pp. 338-345.

[5] Girgensohn, A., Boreczky, J., Wilcox, L., and Foote, J. Facilitating video access by visualizing automatic analysis. *Proceedings of Interact '99*. To appear.

[6] Harrison, B. and Baecker, R. M. Designing video annotation and analysis systems, *Graphics Interface '92*. Morgan-Kaufmann, pp. 157-166.

[7] Isaacs, E. A., Morris, T., and Rodriguez, T.K. A forum for supporting interactive presentations to distributed audiences. *Proceedings of CSCW '94*. ACM Press, pp. 405-416.

[8] Lamming, M. and Newman, W. Activity-based information technology in support of personal memory. Technical Report EPC-1991-103, Rank Xerox, EuroPARC, 1991.

[9] MacKay, W. E. EVA: An experimental video annotator for symbolic analysis of video data, *SIGCHI Bulletin*, 21 (2), 68-71. 1989. ACM Press.

[10] Minneman, S., Harrison, S., Janssen, B., Kurtenbach, G., Moran, T., Smith, I., and van Melle, B. A confederation of tools for capturing and accessing collaborative activity. *Proceedings of the ACM Multimedia '95 Conference*. ACM Press, pp. 523-534.

[11] Moran, T. P., Chiu, P., Harrison, S., Kurtenbach, G., Minneman, S., and van Melle, W. Evolutionary engagement in an ongoing collaborative work process: a case study. *Proceedings of CSCW '96*. ACM Press, pp. 150-159.

[12] Moran, T. P., Palen, L., Harrison, S., Chiu, P., Kimber, D., Minneman, S., van Melle, W., and Zellweger, P. "I'll get that off the audio": a case study of salvaging multimedia meeting records. *Proceedings of CHI '97*. ACM Press, pp. 202-209.

[13] Pedersen, E. R., McCall, K., Moran, T. P., and Halasz, F. G. Tivoli: An electronic whiteboard for informal workgroup meetings. *Proceedings of INTERCHI '93*. ACM Press, pp. 391-398.

[14] Stifelman, L. *The Audio Notebook: Paper and Pen Interaction with Structured Speech*. PhD Thesis. MIT, 1997.

[15] Trigg, R. Computer support for transcribing recorded activity, *SIGCHI Bulletin*, 21 (2), 68-71. 1989. ACM Press.

[16] Weber, K. and Poon, A. Marquee: a tool for real-time video logging. *Proceedings of CHI '94*. ACM Press, pp. 58-64.

[17] Whittaker, S., Hyland, P., and Wiley, M. Filochat: handwritten notes provide access to recorded conversations. *Proceedings of CHI '94*. ACM Press, pp. 271-276.

[18] Wilcox, L. D., Schilit, B. N., and Sawhney, N. Dynomite: A Dynamically Organized Ink and Audio Notebook. *Proceedings of CHI '97*. ACM Press, pp. 186-193.

[19] Wolf, C., Rhyne, J., and Briggs, L. Communication and information retrieval with a pen-based meeting support tool. *Proceedings of CSCW '92*. ACM Press, pp. 322-329.