Supporting An Integrated Paper-Digital Workflow for Observational Research

Nadir Weibel, Adam Fouse, Edwin Hutchins, and James D. Hollan

Distributed Cognition and Human-Computer Interaction Laboratory
Department of Cognitive Science, University of California San Diego
{weibel,afouse,ehutchins,hollan}@ucsd.edu

ABSTRACT

The intertwining of everyday life and computation, along with a new generation of inexpensive digital recording devices and storage facilities, is revolutionizing our ability to collect and analyze human activity data. Such ubiquitous data collection has an exciting potential to augment human cognition and radically improve information-intensive work. In this paper we introduce a system to aid the process of data collection and analysis during observational research by providing non-intrusive automatic capture of paper-based annotations. The system exploits current note-taking practices and incorporates digital pen technology. We describe the development, deployment and use of the system for interactive visualization and annotation of multiple stream of video and other types of time-based data.

Author Keywords

Observational research, Digital pens, Video analysis, Interactive visualization

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation: User Interfaces—*Evaluation/methodology*;
H.5.3 Information Interfaces and Presentation: Group and Organization Interfaces—*Evaluation/methodology*

General Terms

Design, Human Factors, Experimentation, Measurement

INTRODUCTION

A new generation of inexpensive digital recording devices and storage facilities is revolutionizing data collection in behavioral science, extending it into situations that have not typically been accessible and enabling examination of the fine detail of action captured in meaningful settings. Researchers are taking advantage of increasingly inexpensive digital video and storage facilities to assemble extensive data collections of human activity captured in real-world settings. The ability to record and share such data has created a critical moment in the practice and scope of behavioral research. While such data are certainly needed, more data cannot be the whole answer, since many researchers already feel that they are drowning in data. Data without appropriate theoretical and analytical frameworks do not lead to scientific advances. The main obstacles to fully capitalizing on this opportunity are the huge time investment required for analysis.

Because of the complexity and time required for analysis, current methods often result in haphazard and incomplete analyses or, all too commonly, to no analysis at all of much of the data. Even dataset navigation is cumbersome. Data records are chosen for analysis because of recording quality, interesting phenomena, and interaction density, often producing a haphazard sampling of the recorded set.

Good researchers have a nose for good data, but also have a tendency to focus on small segments of the record that contain "interesting" behavior, analyze them intensively, and then move on to the next project. When analysis is so costly, few analyses can be done – so datasets are severely underutilized – and researchers come to have a large investment in the chosen data segments. Since each analysis may appear as an isolated case study, it can be difficult to know how common the observed phenomena may be. Larger patterns and contradictory cases can easily go unnoticed. Well-known human confirmation biases can affect the quality of the science when each analysis requires so much effort.

In the behavioral sciences, researchers transcribe and code data in a wide variety of ways, creating new re-representations of the original events [9]. Currently the coordination of multiple re-representations with the original data is typically done by hand, or not at all. Since this re-representation process – including all sorts of transcription, coding system development and implementation, and re-description – is what allows us to do science [5], even small improvements in automating coding, transcription, or coordination of representations can be crucially important.

No matter what the transcription or the specific coding for the undergoing analysis will look like, researchers still rely on the use of paper-based documents and hand-written notes

both during and after observing the phenomena currently under study. This is largely motivated by the affordances that pen and paper still present even in today's highly digitized work environments. Sellen and Harper [20] explained why the predictions of a paperless office have proven to be unrealistic. They outlined how and why paper still remains the preferred way to support a wide variety of daily activities.

Paper as a medium has many advantages over digital media in terms of how people can work with it, both individually and in groups. It is portable, cheap and robust. It is much more convenient to scan through a book than to browse a digital document. Paper supports forms of collaboration and interaction that are difficult to mimic in current digital worlds. Even with the goal to go paperless, and despite reducing the printing of unneeded paper documents, people still use notebooks or loose paper sheets and a regular pen as instruments for their first informal notes. In observational research, it is a standard procedure to take notes in real-time while recording the studied phenomena and paper documents often represent the only viable way to support the flexibility needed in this situation. These notes commonly provide an important resource for tackling the analysis of recorded events and are used both as a starting point and as an underlying reference throughout the analysis process.

Consider the following scenario:

A research group is studying the coordination of pilots in a commercial aviation cockpit. To better understand the cognitive activity of the pilot and co-pilot, multiple sources of data are recorded during a training session in a full-motion simulator: two angles of video, one from an over-the-shoulder perspective, and one from a camera mounted to the pilot's hat; two channels of audio from label-mounted microphones to capture the individual speech of each pilot; a data stream from the simulator including flight variables such as altitude and airspeed as well as interactions with cockpit controls; notes taken by observers during the simulated flight. As the flight progresses, the researchers take notes about the activity of the pilots, recording errors, communication between pilots, and other notable actions.

After the simulation is over, researchers are faced with the task of analyzing the data. However, there are a number of things that must be done before in-depth analysis can occur. First, the data sources need to be aligned, and then a high-level timeline of the activity needs to be constructed. For example, the video will include incidental activity before and after the specific activity that is the focus of the research, and even within the main activity there is a high-level temporal structure that needs to be understood before more detailed analysis can occur. This may include identifying the phases of flight, and specific maneuvers that the pilots performed. To create this high-level timeline requires going through the video in a linear fashion, and recording the timecodes that correspond to the segmentation points between events or video segments.

Once this high-level structure is created, then individual moments need to be identified that represent aspects of the activity that deserve thorough analysis. This often requires spending additional time watching the video, but also can be aided by using the notes taken during the activity. These notes may include timecodes and descriptions of points in the video that correspond to particularly interesting episodes. To view a moment in the video, the timecode may need to be translated to the timecode system of the video (such as from a time of day to a relative time in the video) before moving to the segment of the video that corresponds to the notes.

In this paper we introduce a system that supports observational research by linking paper-based notes with video recordings. We demonstrate how current digital pen technologies can be integrated within the standard workflow of observational researchers and how paper-based information supports interactive visualizations and analysis of data. We start by outlining current practices for note-taking and video analysis and highlighting technologies for augmenting paper documents. We then describe our system, and highlight its role both during the initial observation and as part of subsequent analysis. We then discuss how the system was deployed and used in specific situations and discuss future work.

RELATED WORK
Observational research increasingly involves fine-grained video analysis of activity. Researchers are able to record details of human activity for subsequent off-line analysis and processing. Nevertheless, the analysis process typically begins in the form of activities that researchers undertake in parallel with observation. While observing, researchers may identify the key aspects of the observed phenomena or perhaps even prepare a mental plan for the analysis. As pointed out by Schön [19], *reflection* is an important process that allows people to better structure their analysis and occurs both during the observation activity – *reflection-in-action* – and after the fact, while looking back at the observed activities and during the analysis process – *reflection-on-action*. In this section we present existing approaches supporting both activities, specifically highlighting how the process of note-taking is particularly important during the on-line reflection, and how digital video analysis tools help in the off-line reflection. We also introduce hybrid paper-digital systems, tools and technologies that are showing great potential for enabling the interconnection of these two activities.

Paper-based notes
For obvious environmental issues, there is a widespread interest in *going paperless*, exploiting more and more digital-only tools for everyday activities. A range of tools and new systems such as Tablet PCs, PDAs, and iPads seem to open the way for this transition and they effectively enable shifting some paper-based activities directly to the digital world. However, as documented by many studies [1, 10, 13], paper use continues to grow. The process of informal or semi-structured note taking is still heavily paper-based [20]. Practices such as arranging papers, post-it notes or scraps in space which were formerly taken as indices of

disorganization are now recognized as information management techniques [2]. In particular, personal notebooks let people save, filter and transform information, not by supplementing human memory but rather helping users select and synthesize the information they handle [23]. Moreover, the physical layout and spatial organization of paper-based notes may carry information about the ongoing activity and is rarely supported by digital tools alone [11]. It is not only the content of a medium that carries information but also the way people handle and organize artifacts within their environment. The transition to the sole use of digital aids would also mean losing aspects of paper notebooks such as durability or ease of reading [15] that are particularly important in many daily activities or specific scientific settings such as observational research. Studies show how researchers use different notebooks depending on the expected life time of the information they save [24]. In the biology setting, for example, paper laboratory notebooks are used for the long term storage of information.

Even if paper presents many affordances, it still does not always exploit the interaction level which is achieved by digital documents. Therefore, the vision of more interactive documents, supporting researchers in the management of different kinds of digital information, could not be achieved by using paper documents alone. We conjecture that the optimal solution lies in between, where "paper and electronic document tools work in concert and organizational processes make optimal use of both" [20]. Digitally augmented paper documents represent one example of the required evolution from a traditional paper technology into a hybrid system combining paper and digital documents.

Allowing a standard paper document to be linked with a digital system has been investigated in numerous research projects in the past. An early example is Wellner's Digital Desk [26], a physical desk augmented with digital capabilities that allowed the tracking of hand-written information on paper by means of a ceiling-mounted camera. In the setting of paper-based notes, Mackay et al. [12] investigated how laboratory notebooks could be augmented with digital aids, by capturing hand-written information and linking it to searchable electronic versions of the notebook. The Audio Notebook [22] adds audio recording to a normal notebook, allowing paper-based notes to be automatically linked to speech recording, while the A-book [12] enables interactive visualization to be added on top of a paper notebook by using a PDA as a "physical information lens" providing bidirectional links between paper and digital information.

Other projects addressed field observations in mobile environments. ButterflyNet [27] is a mobile capture and access system supporting pen and paper interactions for field scientists. It supports the storage of handwritten annotations and their combination with heterogeneous data (GPS track logs, pictures or sensor readings). The paper-based annotations can be used as a browser to access data stored within the ButterflyNet repository by tapping an interactive notebook page. Whereas ButterflyNet focuses on heterogeneous field-data capture, Prism [24] is a hybrid paper-digital notebook

integrating information from the desktop and the Web with paper notebooks. Other systems go beyond linking paper-based information to the digital world to allowing changes in one media to produce changes in another. ProofRite [3] and PaperProof [25] allow users to edit and annotate paper documents and have the results accomplished in the digital version. This enables information to freely flow from one representation to the other.

While technologies enabling hybrid paper-digital systems mainly support the real-time collection of data, researchers typically confront the complex task of linking their paper-based notes to related sections of video in increasingly large video data archives. In the next section we describe emerging video analysis practices and then discuss how those practices can be better supported.

Video Analysis Practice

In response to the growing use of video as a research tool, a number of researchers have identified the elements of a video research workflow, and how these activities should be supported by software tools. Pea and Hoffert [16] proposed a research process for working with digital video. The description of this workflow process included all of the stages from planning to publication, and gave detailed focus to an iterative process of analysis. Pea and Lemke [17] made further mention of the iterative nature of video research, noting the need to "chunk" the video records into segments that may be defined in a number of ways.

While there are a range of activities involved in video research, the task of video annotation has received particular focus, both from researchers and software developers. The requirements for the task of video annotation have also been studied to be able to better design and compare software tools. Hagedorn, Hailpern, and Karahalios [7] drew on interviews with experienced video researchers to identify a set of requirements for effective video annotation software based on current practices. Among these requirements was the need to facilitate coding workflow, capturing and displaying appropriate types of additional data, and flexible forms of video playback. Identification of these requirements was driven by the development of the *VCode* and *VData* pair of video analysis tools.

Hofmann, Hollender, and Fellner [8] also established a set of capabilities needed to support the video research process. The general categories of tasks that were identified are *configuration*, *segmentation*, *annotation*, *exploration*, and *externalization*. Among the specific requirements they identified, the ones particularly germane to the current discussion are supporting the annotation style of individual users, synchronization of data sources, and supporting exploration by providing multiple ways of navigating the data. This set of requirements was designed to be used as a comparison across a range of video annotations tools, including *ELAN*, *The Observer*, *Transana*, and *WebDIVER*, among others.

INTERACTIVE PAPER

In the previous section we presented a range of approaches that combine paper-based notes with digital media and present them together as part of the same interface. The *Anoto Digital Pen and Paper technology*[1] enables paper to be turned into an interactive medium, creating *Interactive Paper* documents.

Paper documents can function as interactive interfaces that recognize handwritten information, link to digital media such as sound or voice recordings, or trigger specific digital applications on a separate computer either in real time (by streaming the collected information over bluetooth) or in batch-mode (by connecting the digital pen to a computing device). This technology enables users to exploit rich digital services while keeping the natural interaction common in traditional pen and paper interfaces. The system is based on a special digital pen that integrates a processor, memory, and an infrared camera able to interpret a unique dot-pattern printed on standard paper. By decoding the printed pattern, the device can track the pen's position on paper in real-time.

Anoto provides tools to help designers and developers build form processing applications. In order to allow a more flexible access to Anoto-based documents, as well as a more general approach towards the management of interactive paper applications, different approaches have been introduced in recent years. The Paper Augmented Digital Documents infrastructure (PADD) [6], developed at the University of Maryland, supports the basic document management across the paper-digital boundary. More advanced support for the management of complex interactive paper applications was introduced at ETH Zurich by the iPaper/iServer framework [14] and at Stanford University by PaperToolkit [28]. Designers and developers working with these frameworks are no longer restricted to form-based applications, but they are able to create and manage a range of innovative interactive applications.

Recently, a novel kind of Anoto digital pen, the Livescribe Pulse Smartpen[2] was introduced. This pen, shown in Figure 1, integrates a computer, speakers, a microphone and an OLED display. While the main application of Livescribe Pulse pens consists of recording lectures or meetings and synchronizing them with handwritten notes, Livescribe has released SDKs that allow developers and designers to create multimodal applications and deploy them as stand-alone programs on the pen itself.

SYSTEM DESCRIPTION

By exploiting the capabilities of interactive paper technology, we developed a system to ease some of the tedious tasks associated with video analysis. Again consider the scenario presented in the introduction, studying pilot coordination in a simulated flight cockpit. The process of analyzing the video data involves a variety of often tedious activities:

1. *Visual scanning to identify moments of activity.*
   Most video analysis tasks require visual search. As the video or videos are played back, potentially at different playback rates, the researcher watches the video to identify the types of activity that are related to the research focus.

2. *Finding points that match descriptions in the notes.*
   Notes recorded during initial observation serve as a guide to select episodes for further analysis, but these episodes need to be found in the video. If specific times were not recorded, then the descriptions in the notes need to be matched to activity being viewed in the videos. While the linearly sequential nature of the video may help in this task if the notes are chronologically ordered, it still remains a tedious and time consuming task.

3. *Finding points to match written times.*
   If times were recorded, they may not be the same as the time codes that are present with the video. For example, times may be recorded as a time of day, which would then need to be converted to a time code in the video.

4. *Recording annotations and video time-codes.*
   Once moments in the video are found, a record needs to be made of the time-codes and corresponding descriptions.

Now imagine the following addition to the initial scenario:

During initial observation of the simulated flight, the researchers record their observations with a digital pen on interactive paper. One member of the team, unfamiliar with the system, takes notes as he normally does with pen and paper, while another, more experienced, member makes marks to indicate specific types of activity that she knows will need to be analyzes further, and to code particular actions by each pilot. When the researchers get back to the lab, they dock their pens, and transfer the video to a computer. They are then able to view the video with an automatically generated interactive timeline that shows the time points and categories of the special marks made by the second researcher. By clicking on the timeline, they are able to instantly go to the corresponding points in the videos. In addition, they are able to see the notes made by the first researcher, and clicking on them allows the researchers to access specific points in the videos.
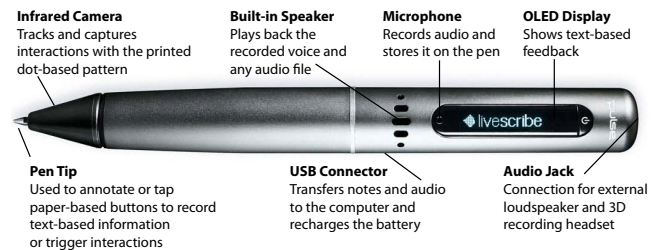

Infrared Camera
Tracks and captures interactions with the printed dot-based pattern

Built-in Speaker
Plays back the recorded voice and any audio file

Microphone
Records audio and stores it on the pen

OLED Display
Shows text-based feedback

Pen Tip
Used to annotate or tap paper-based buttons to record text-based information or trigger interactions

USB Connector
Transfers notes and audio to the computer and recharges the battery

Audio Jack
Connection for external loudspeaker and 3D recording headset

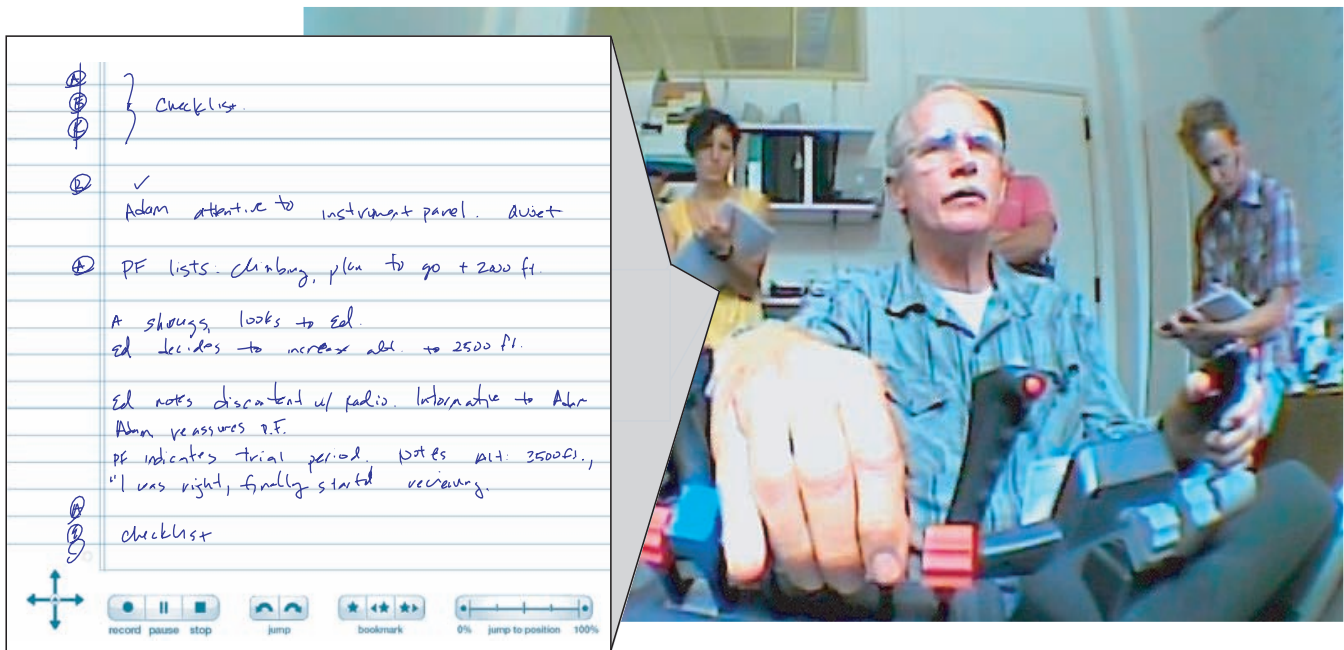**Figure 1. Livescribe Pulse Smart Pen**

Figure 2. Initial observation session, showing the observed subject and three researchers (right), and the paper notes of one of the researchers (left)

Our goal in designing the system was to facilitate and simplify analysis activities by taking advantage of interactive paper technology. This allows integration of an interactive visualization system, natural note-taking practice, and real-time coding of data. Note that while this scenario focused on flight research, we have tested components of this system in a number of domains, including research focused on responder interaction during emergencies[3], children in a Gamelan musical ensemble[4], and assistive technology for speech therapy [18]. The types of activity supported by this system can be applied to video-based observational research in many domains. In all cases, there is a need for better ways of navigating the video during analysis and taking advantage of researchers' efforts during observation.

Support during initial observation

During initial observation, paper notes are taken using a digital pen. In the current system, we use Livescribe Pulse pens. Support for this phase is designed to be a natural extension of normal note-taking, but also to take advantage of the capabilities of the digital pen. At the most basic level, notes can be taken as if the researcher was using a non-digital pen. The system makes minimal demands on the researcher, but still takes advantage during analysis of the digital nature of the notes. At a more advanced level, custom gestures can be defined that can be automatically translated to more semantically meaningful annotations of the data. Figure 2 illustrates a typical observation session, with the subject being studied in the foreground and three researchers observing the scene in the background. The subject is performing a flight simulation and two of the researchers are taking notes on the specific maneuvers he is performing. The notes of the researcher on the left-hand side of the figure are highlighted.

_____
[3]http://www.wiisard.org
[4]http://tdlc.ucsd.edu/research/highlights/rh-gamelan.html

While the described observation session might seem just like any other ethnographic-style activity, the usage of the Livescribe digital pen enables the paper-based notes to be directly linked with all other digital recording and data collected during the specific session. The pen is equipped with a dedicated software component called *Ethnographer++* that has been installed on the pen and runs during the whole session. Ethnographer++ records every pen stroke and is able to timestamp them in such a way that they can be linked back to any other time-based data, for instance the videos of the activity or specific sensor data such as GPS or simulator logs. Additionally, if required, Ethnographer++ may also record the surrounding audio exploiting the microphone integrated in the Livescribe pen. The recorded audio can also be linked to any of the other data by means of an additional timestamp. Finally, Ethnographer++ supports the definition of special regions on the paper sheets that can react in different ways to pen-based interaction, providing specific audio/visual feedback.

In order to collect the information needed to link the time-stamped hand-writing and the recorded audio, three different set of files are recorded on the Livescribe pen by Ethnographer++: (1) a set of XML files representing paper-based strokes, (2) a set of audio files natively stored by the Livescribe pen as AAC (Advanced Audio Coding) files, and (3) additional XML files encoding synchronization information between the paper notes and the recorded audio. For every new Ethnographer++ session, a new dedicated folder is created in the pen file-system and all data regarding that specific session are stored in it. The paper-based strokes are collected in real-time as the user writes on paper, and temporary stored as a `Note` object within the application. Ethnographer++ notes are a variant of interactive paper notes defined as part of the iPaper framework briefly introduced in the previous section: a `Note` consists of a set of `Stroke` objects defined

by time-stamped (x,y) `Point` objects. `Stroke` objects are collected every time a *pen-up* event is detected (when users lift the pen) and added to the `Note` objects, which additionally encode the start time of the paper-based note and the page number where the note was collected from. Notes are serialized as XML files as soon as users interact with a new page, after a specific timeout, and before the Ethnographer++ application is terminated. Figure 3(a) shows an example of a note recorded as XML, while Figure 3(b), highlights the synchronization information, encoding the start and stop time of every single audio file associated with the current session.

As we already highlighted earlier, the introduction of Ethnographer++ and a Livescribe pen are totally transparent with respect to the observation session. This allows researchers to focus on their observations and use the paper-based notes as bookmarks or reminders of specific events as they usually do. The left hand-side of Figure 2 illustrates such an example. The figure shows how the researcher added specific symbols and bookmarks on the side of the paper notes to mark special events. We will see in the following sections how these special marks are actually a powerful tool for indexing the recorded session in real-time.

Data consolidation

Once the observation task is over, researchers have to retrieve data from the different recording devices and consolidate it in order to start analysis. Video data can be directly transferred from the internal memory of the camera and simulator logs can be retrieved from the computers, while information on paper-based notes and possibly recorded audio has to be extracted from the Livescribe pen.

In order to extract data from it, the pen has to be connected over USB to a computer, where *PenDesktop*, a dedicated software built on top of the Livescribe Desktop SDK[5], is able to query the pen internal storage and persistently save the needed information on the computer.

After being transmitted to the local computer, the pen data is analyzed and *pen sessions* are automatically created for every recorded observation. The PenDesktop software extracts the single audio and note files and stores them in separate session folders. While PenDesktop provides a front-end interface for managing the recorded sessions, enabling direct access to the recorded files and supporting session renaming and merging, we decided to base the actual storage of

```
<note startTime="1284017295"       <AudioSync
page="2426622162526260">             recording="1284017295.aac">
  <stroke>                            <start>1284017295</start>
    <point>                           <stop>1284019141</stop>
      <x>259.3875</x>               </AudioSync>
      <y>21.675</y>
      <time>1284017557</time>
    </point>
    ...
  </stroke>
  ...
</note>
```

  (a) Note object   (b) Synchronization object

Figure 3. XML representations

the single files on a hierarchical file-based structure. This allows users to easily exchange specific data (e.g. a single audio recording) and get access to the single files without the need of interacting with the PenDesktop user interface.

After retrieving the raw data from the pen and creating structured sessions, these are automatically synchronized with a central server. A distributed repository (currently based on DropBox[6]) is shared by different users and mounted on the central server component. The role of the server component is to perform automatic analysis of the recorded pen-based and audio data and opportunistically provide users with enhanced data about specific observations or tasks. The server is continuously watching the pen session root directory for new data, and as soon as new pen data is made available by any of the users through their local DropBox folders, it starts one or more analyses based on specific session requirements, eventually generating rich-data sets to be used for further analysis. The integration of the DropBox mechanism also enables users on the local computers to be automatically notified as soon as the server component finishes analyzing the data, and to get direct access to the data on their local computer through the shared repository.

The server component of PenDesktop can be instructed to potentially execute any kind of analysis on the raw pen data, depending on the specific task and observation. However, we believe that some specific analyses can be applied to most pen and paper notes and we enable them by default. Currently we have implemented a handwriting recognition analysis plugin for the server based on the MyScript Intelligent Character Recognition (ICR) software[7] and a gesture recognition analysis plugin based on the iGesture framework [21]. These two plugins analyze different pen-based data such as the special symbols on the margin of the notes showed in Figure 2 or the hand-written notes shown in the rest of the page, and generate rich XML data encoding digital representations of the gesture and hand-written text that have been recorded during a session. These XML files can be integrated and associated to represent gesture+text interactions that are meaningful for the specific ethnographic sessions and are made available to the researcher automatically just seconds after plugging the Livescribe pen to the computer.

Support during analysis

Once the data from the pen has been transferred to a desktop computer and processed, it can be used to aid analysis through a tool that we have developed called ChronoViz[8], which supports visualization and analysis of multiple streams of multimodal, time-based data [4]. ChronoViz allows researchers to visualize time-based data from multiple sources and manually or automatically annotate the data. The data sources can include multiple video files, audio files, log files (encoded in a variety of formats), and transcript data. The main ChronoViz interface consists of a video pane, which can show multiple videos, and one or more timelines
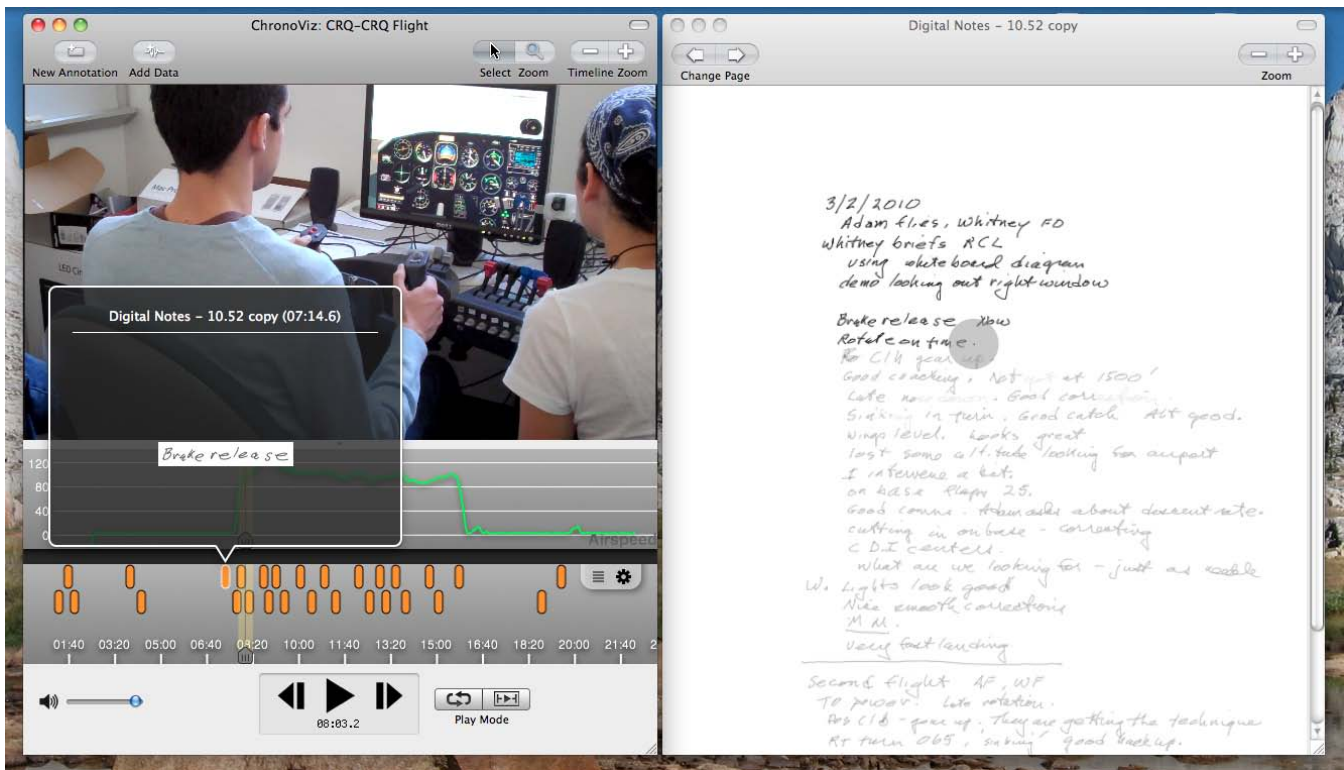

Figure 4. ChronoViz analysis, showing digital notes as annotations on a timeline (left) and a digital representation of the paper page (right).

arrayed below the video. A single timeline is shown by default, but multiple timelines can be added to visualize additional data, group annotations, or view different time scales. In the example shown on the left side of Figure 4, there are two timelines, one showing annotations and one showing a time-series graph of airspeed from a simulated flight.

Data annotations in ChronoViz can be visualized on their own timelines, shown in lists, or overlaid on other data types. For example, a set of annotations might be created in relation to moments in a video that indicate particular types of activity. The corresponding segments of time can then be highlighted on a time series graph to depict how various variable values changed during those aspects of the activity. The

Figure 5. Paper-based notes: original (left), interactive reprint (right).

annotations are color-coded with researcher-defined colors, and can be shown on the timelines and on maps. More detail about the annotations on the timeline can be shown by hovering the mouse cursor over the annotation marker, resulting in a pop-up detail window. By combining the paper-based digital notes recorded during observation of the activity with the other data that was collected, annotation and navigation of these diverse sources of data can be improved. After importing the pen sessions data generated by the PenDesktop tool described in the previous section, the digital pen data can be integrated in a number of ways. First, the notes as a whole can be displayed as a digital version of the paper notes, as shown in Figure 4. The pen marks are initially displayed in gray, but as the researcher moves through the data by interacting with any of the other data streams, the letters get darkened to indicate which notes correspond to the current time in the data. This synchronization also works in the opposite direction: the researcher can change the current time position in the data (e.g., the frame shown in the video) by clicking on particular notes to jump to the corresponding points in the rest of the data.
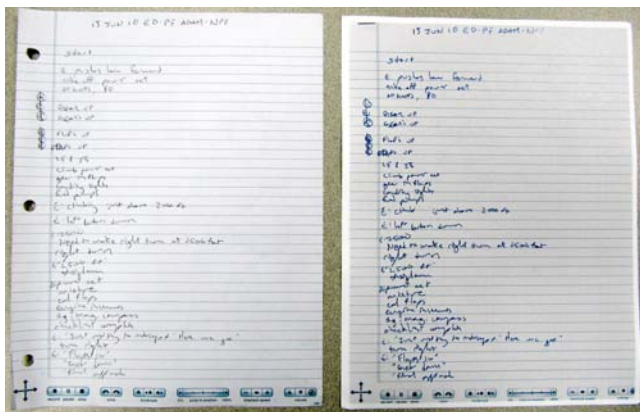
Apart from creating the pen session data exploited directly by ChronoViz, the PenDesktop tool is also able to create an interactive version of the paper-based notes that can be used as a direct interface to ChronoViz, in conjunction with a bluetooth Anoto pen that communicates with the computer in streaming mode. Because of incompatibilities of Anoto pattern between Livescribe applications and Anoto streaming applications, we currently have to reprint the collected strokes on Anoto "streaming pattern". A special PenDesktop

print plugin retrieves all strokes collected for a specific session and generates a digital representation of the paper notes as a PDF document. This document is then automatically augmented with streaming pattern and sent to the printer. As illustrated in Figure 5, this new document represents an interactive copy that is nearly identical to the original notes.

ChronoViz supports a custom URI scheme that allows researchers to issue specific requests to the visualization software based on the position of paper interaction. We use the iPaper framework outlined previously to communicate with a wireless bluetooth Anoto DP-201 pen and to extract information about the specific paper position. This position is then encoded in a special URI request which is finally sent to ChronoViz.

```
chronoviz://note?page=2426622162526260&x=27&y=112
```

This page and position information can be interpreted in the context of pen session data that has already been loaded into ChronoViz and aligned with the rest of the data to generate a timecode and match the displayed time of the videos to the time at which the note was originally written. By taking advantage of the actual paper form of the notes and exploiting the connection between ChronoViz and the interactive paper add-on, the researcher can use the notes to navigate video in a way very similar to using the on-screen representation of notes. The researcher can use the pen to tap on a particular note on paper, and the on-screen representation of the rest of the data will move to the corresponding time point.

Finally, in addition to interaction with the entire page of notes in digital or paper form, markers can appear on a timeline below the video that correspond to the time that a particular note was written. These annotations can be generated in several ways. First, they are generated by temporal segmentation. Essentially, pauses in the act of taking notes corresponds to the start of a new annotation. When the researcher hovers over the marker, an image of the note appears. These markers can also be clicked to move to that time, and provide a visualization of the pattern of note taking. For example, if a large number of markers are clustered in one location on the timeline, it means that a lot of notes were taken around that time, and this may be a clear visual indication of interesting activity. Additionally, if multiple sets of notes have been loaded, the markers are color-coded to group each person's notes, and the alignment of different observers' notes may also be an indication of interesting moments of activity.

In addition to temporal segmentation, if pen gestures have been used when writing the original notes, then more semantically meaningful annotations can be created that correspond to particular gestures or groups of gestures. Basic gestures can be used to identify individual moments, and may be as simple as a mark to indicate a time when something interesting happened or may be more complex sets of gestures to code particular classes of activity. These different classes would then appear as different sets of color-coded annotations. Finally, gestures can be linked together to code regions of activity. These linked gestures would then be used to create *duration* annotations that appear as bars on the timeline with endpoints that correspond to the start and end times of the activity. Through the use of pen gestures, rich coding can be created in real-time during observation and be immediately available as an entry to further analysis.

DISCUSSION

Paper-based notes are an important resource that are commonly used for structuring subsequent analysis. The introduction of a natural way to capture notes and integrate them automatically within a researcher's analysis workflow, without interfering with the established observation and analysis processes, creates novel and interesting possibilities for interactive analysis of the observed phenomena.

In general, recording notes with interactive paper technology has the appealing property of supporting different ways of working. Notes can be recorded with no change in practice, as if using traditional pen and paper, or researchers can make use of more advanced pen gestures. We see a range of advantages that the integration of this technology with interactive visualization offers over traditional methods.

Time efficiency

One of the goals of many video annotation systems is to decrease the time required for analysis. As mentioned in the introduction, many researchers who use video-based observational methods feel that they simply have too much data and not enough time to properly analyze it. Our system addresses this issue in part through digital interaction with notes in alignment with video and other data sources. Points of interest that were identified during initial observation can be found with little effort, by simply clicking on a note to get to a point in the video. This has the effect of both making the notes more valuable, by empowering the researcher to make more efficient use of them during analysis, and making navigation of the video much easier, by exploiting effort already expended during initial observation.

Real-time personal bookmarks

Navigation of video can also be improved through the use of real-time personal bookmarks, created by pen gestures. These gestures can be systematized, which can be beneficial to collaborative analysis and to computational interpretation of the gestures, but they can also be more abstract. One characteristic of the personally hand-written notes is that they are still useful even if they are meaningful only to the person who wrote the note. For example, while an abstract mark in the margin might not be significant to a second observer, the original note-taker will likely be able to remember some aspect of the activity that the mark corresponds to, or the reason for making the mark. It could simply be an indication that something interesting happened at that time and is worth analyzing in more detail. This type of note taking is possible because of the integration of the notes with the rest of the data. During analysis, marks in the margin can be clicked on-screen or tapped on paper to go directly to that time in the video, without needing any additional information.

In addition, we predict that integration of digital notes with analysis may shift practices with regard to how notes are taken during observation. For example, since the pen records the time for each stroke, and these times can be aligned with video of the data, there is much less need to record indexical information, such as time codes or references to the chronology of activity. Even the common chronological nature of notes (i.e., where the order of notes on the page corresponds to the order of activity) could be changed in favor of spatial groupings of related notes.

Augmented analysis

More systematic and semantically meaningful pen gestures offer additional benefits. While personally relevant marks can function as "personal bookmarks," gestures that have been given a defined meaning within the system can be used to generate initial coding of the data in real-time. Through the use of different marks, a timeline can be automatically constructed that shows an overview, using color-coded markers to visually identify different types of activity. This visual structure can then serve as the basis for further investigation, offering a rich starting point for further analysis.

For example, in Figure 2, the sheet of notes includes a series of circles letters in the margin. These are part of a coding scheme that has been designed for observation of the pilot/co-pilot coordination in a cockpit. In this case, an "A" is recorded when the pilot asks the co-pilot for a piece of information or to perform an action, a "E" is recorded when the co-pilot executes a requested action, and an "R" is recorded when the co-pilot gives the pilot a verbal response. Upon completion of one of these instances of pilot/co-pilot interaction, a line can be drawn through the letters, which results in the individual actions being combined into a single *duration* annotation when the pen data is brought into ChronoViz.

Notably, this is a flexible system, in which gestures can be defined that are meaningful to the researchers, or that are specific to the activity that is being observed. This can range from using particular letters in the margin to identify different actors in the event or types of action, to more comprehensive codings of activity using specially designed forms. Since the gestures can be defined with respect to both the form of the pen stroke as well as the location on the page, different types of note-taking can be combined on the same page, such as having an area for normal free-text note taking and an area for more structured coding. The current system allows arbitrary gestures to be defined and interpreted at synchronization time. By combining hand-writing recognition and different gesture subsets, users can define their own composite gestures and assign them to specific activities. The iGesture library currently integrated in the system natively supports a number of pre-defined gestures (Palm-based graffiti numbers and letters and Microsoft application gestures[9]), but it also allows users to create their own gesture sets through a brief training phase. Predefined gestures already known to the application can be directly assigned to activities, while new gestures defined in real-time first need to be linked to a particular type of activity.

---

[9]http://msdn2.microsoft.com/en-us/library/ms704830.aspx

Information interaction

While on-screen integration of notes offers the advantages already listed, interaction with physical paper notes during analysis has several additional advantages. First, display space can be more effectively used; since there no longer needs to be a digital representation on screen, other data can be shown or existing data can be shown larger. Second, the flipping between pages of notes seems to be more natural with sheets of paper than clicking buttons on screen. Finally, this opens the possibility for better collaborative analysis. Multiple researchers can have their own notes, which can each be used to control movement in the data. This can be particularly useful when several researchers are looking at data projected onto a screen and the individual researchers with their own notes and a bluetooth Anoto pen can direct video to the specific point that they are discussing.

CONCLUSION AND FUTURE WORK

In this paper we have presented a system to support video-based observational research by creating an integrated paper-digital workflow. This allows notes recorded during observation of activity to serve as an interactive basis for later data navigation and analysis. We believe this has the potential to increase the efficacy of video as a scientific research tool by facilitating annotation and analysis. Since the existing benefits provided by this system are promising, we have identified several extensions of the paper-digital workflow supported by the system that we intend to pursue.

First, we plan to more extensively investigate how existing patterns of writing may be exploited to recognize more information in written notes. The current system supports a flexible way of recognizing predefined gestures, but there are a number of extensions that we would like to consider. One type of pen usage involves forms. For example, ethological researchers often use forms to complete ethograms of animal behavior. Integrating the structured data collected from these forms directly in the video analysis tool as coded annotations would speed data analysis. Another example of preprinted forms are diagrams or charts of intended activity, such as an altitude profile to be followed during a flight training session. Enabling researchers to make marks in relation the preprinted profile, and having the system automatically interpret these marks in the context of the preprinted information, would be an important advance.

While interpreting marks in relation to preprinted structures would be quite useful, there is often structure in the notes themselves that may be exploited as well. The spatial organization of notes may be used to infer categories, while drawn elements such as lines and boxes may have meaning that could be beneficial, beyond a simple recognition of the stroke-based gestures. Structures such as hand-drawn tables or graphs could be recognized and integrated with the data in the same way as the rest of the notes. Additionally, explicit ways to support quantization of activity such as recognizing tally marks and interpreting them as numeric annotations would be useful.

Finally, in addition to supporting different types of handwritten notes, another possible extension could support other interactive uses of the pen. For example, in our discussions with several researchers, we found that a common activity for high-level analysis of video is to watch the video on a computer screen while writing notes on paper. The current system already provides basic support for this activity if the video is played straight through, but we have planned further support by creating a richer connection between the pen activity and the on-screen data. The pen could be used for types of playback control other than jumping to specific times, such as providing paper buttons for adjusting playback rate or video looping behavior. In addition, notes that are written on paper could be integrated in real-time, appearing in the on-screen notes or as annotations.

We are also planning a formal evaluation to further understand the initial promise shown by the existing and planned capabilities. To assist this evaluation, the ChronoViz application can collect data about its use by participant researchers, including temporal patterns of data navigation and annotation. Since researchers can use ChronoViz for data annotation without taking advantage of the paper-digital workflow, we will be able to formally compare analysis activity that takes advantage of these capabilities with analysis activity that does not.

REFERENCES

1. E. Auramäk, M. Robinson, A. Aaltonen, M. Kovalainen, A. Liinamaa, and T. Tuuna-Väiskä. Paperwork at 78kph. In *Proc. CSCW '96*, pages 370–379.
2. M. Bernstein, M. Van Kleek, D. Karger, and M. Schraefel. Information Scraps: How and Why Information Eludes our Personal Information Management Tools. *ACM Transactions on Information Systems (TOIS)*, 26(4):1–46, 2008.
3. K. Conroy, D. Levin, and F. Guimbretière. ProofRite: A Paper-Augmented Word Processor. Technical Report HCIL-2004-22, CS-TR-4652, Human-Computer Interaction Lab, University of Maryland, USA, 2004.
4. A. Fouse and J. D. Hollan. DataPrism: A Tool for Visualizing Multimodal Data. In *Proc. Measuring Behavior 2010*, pages 181–184.
5. C. Goodwin. Professional vision. *American anthropologist*, 96(3):606–633, 1994.
6. F. Guimbretière. Paper Augmented Digital Documents. In *Proc. UIST '03*, pages 51–60.
7. J. Hagedorn, J. Hailpern, and K. Karahalios. VCode and VData: Illustrating a New Framework for Supporting the Video Annotation Workflow. In *Proc. AVI 2008*, pages 317–321.
8. C. Hofmann, N. Hollender, and D. Fellner. Task-and Process-related Design of Video Annotation Systems. In *Proc. Mensch & Computer '09*, pages 173–182.
9. E. Hutchins. *Cognition in the wild*. MIT Press, 1995.
10. Z. Liu and D. Stork. Is Paperless Really More? *Communications of the ACM*, 43(11):94–97, 2000.
11. W. Mackay, A. Fayard, L. Frobert, and L. Médini. Reinventing the Familiar: Exploring an Augmented Reality Design Space for Air Traffic Control. In *Proc. CHI '98*, pages 558–565.
12. W. E. Mackay, G. Pothier, C. Letondal, K. Bøegh, and H. E. Sørensen. The Missing Link: Augmenting Biology Laboratory Notebooks. In *Proc. UIST '02*, pages 41–50, Paris, France.
13. G. Moore and J. O'Hare. Paper's future role. *Paper 360°*, pages 16–18, Apr 2008.
14. M. C. Norrie, B. Signer, and N. Weibel. General Framework for the Rapid Development of Interactive Paper Applications. In *Proc. CoPADD '06*, pages 9–12.
15. K. O'Hara and A. Sellen. A Comparison of Reading Paper and On-line Documents. In *Proc. CHI '97*, pages 335–342.
16. R. Pea and E. Hoffert. Video Workflow in the Learning Sciences: Prospects of Emerging Technologies for Augmenting Work Practices. In R. Goldman, R. Pea, B. Barron, and S. Derry, editors, *Video Research in the Learning Sciences*, pages 427–460. 2007.
17. R. Pea and J. Lemke. Sharing and Reporting Video Work. In S. Derry, editor, *Guidelines for Video Research in Education: Recommendations from an Expert Panel*. Data Research and Development Center, 2007.
18. A. M. Piper, N. Weibel, and J. D. Hollan. Introducing Multimodal Paper-Digital Interfaces for Speech-Language Therapy. In *Proc. ASSETS '10*, pages 203–210.
19. D. Schön. *The reflective practitioner: How professionals think in action*. Basic Books, 1983.
20. A. Sellen and R. Harper. *The Myth of the Paperless Office*. MIT Press, 2003.
21. B. Signer, U. Kurmann, and M. C. Norrie. iGesture: A General Gesture Recognition Framework. In *Proc. ICDAR '07*, pages 954–958.
22. L. Stifelman, B. Arons, and C. Schmandt. The Audio Notebook: Paper and Pen Interaction with Structured Speech. In *Proc. CHI '01*, pages 182–189.
23. A. Tabard. *Supporting Lightweight reflection on personal information*. PhD thesis, Université Paris-Sud & INRIA, 2009.
24. A. Tabard, W. Mackay, and E. Eastmond. From Individual to Collaborative: The evolution of Prism, a Hybrid Laboratory Notebook. In *Proc. CSCW '08*, pages 569–578.
25. N. Weibel, A. Ispas, B. Signer, and M. C. Norrie. PaperProof: A Paper-Digital Proof-Editing System. In *Proc. CHI '08*, pages 2349–2354.
26. P. Wellner. Interacting with Paper on the DigitalDesk. *Communications of the ACM*, 36(7):87–96, Jul 1993.
27. R. B. Yeh, C. Liao, S. R. Klemmer, F. Guimbretière, B. Lee, B. Kakaradov, J. Stamberger, and A. Paepcke. ButterflyNet: A Mobile Capture and Access System for Field Biology Research. In *Proc. CHI '06*, pages 571–580.
28. R. B. Yeh, A. Paepcke, and S. R. Klemmer. Iterative Design and Evaluation of an Event Architecture for Pen and Paper Interfaces. In *Proc. UIST '08*, pages 111–120.