DataPrism: a Tool for Visualizing Multimodal Data

Adam S. Fouse
afouse@cogsci.ucsd.edu

James D. Hollan
hollan@cogsci.ucsd.edu

Distributed Cognition and Human-Computer Interaction Laboratory
Department of Cognitive Science, University of California, San Diego
9500 Gilman Drive, La Jolla, CA, USA, 92093

ABSTRACT
We present DataPrism, a new interactive visualization tool to aid analysis of multimodal activity data. DataPrism enables analysts to visualize, annotate, and link multiple time-based data streams, including video, log files, and paper-based digital notes. Automated analysis is supported through a plug-in mechanism. A usage scenario is employed to describe DataPrism's current facilities.

INTRODUCTION
Advances in technology continue to improve the ability of researchers to study real-world behavior. A new generation of inexpensive digital recording devices and storage facilities are revolutionizing data collection, extending it into situations that have not typically been accessible and enabling examination of the fine detail of action captured in meaningful settings. This is important because in order to understand the dynamics of human activity we must understand its full context, and that can only be accomplished by recording and analyzing data of real-world behavior.

Consider three examples: affordable high-definition video and inexpensive digital storage have enabled recording video of activities from multiple angles; small wearable sensors have made it possible to capture detailed records of complex human behavior; and interaction with digital systems can produce time-stamped logs of performance. Each presents new opportunities to analyze activity in more systematic and precise ways. More generally, the ability to collect these and other forms of rich data provides an unprecedented opportunity for scientific scrutiny of activity as it occurs in real-world everyday situations.

Along with the exciting scientific possibilities created by new data collection abilities there are also clear analysis challenges due to the scale and richness of data that is now readily collected. For example, coding and careful analysis of video data remains extremely time-consuming. Other forms of data, such as sensor readings, while useful as separate data sources, may provide more valuable insights about activity when viewed and analyzed in combination. New visualization tools are needed to address the challenge of analyzing and synthesizing information from multiple data sources and at multiple time scales.

We have developed DataPrism, a tool that allows researchers to visualize time-based data from multiple sources and manually or automatically annotate the data. To elucidate its facilities, we present a scenario from a current use of DataPrism, involving analysis of the behavior of airplane pilots during simulated flight. Understanding their cognitive activity requires analysis of details of pilot behavior that can only be acquired through the collection of a rich set of multimodal data. While we focus here on a specific scenario, DataPrism is designed to support analysis of a wide range of multimodal time-based data. For example, it is currently being used in several other contexts, including analysis of children interacting with museum exhibits and of personal activity data captured with the Microsoft SenseCam [2].

Although several other systems have been developed to support various aspects of this visualization challenge, DataPrism represents a unique focus on the visualization of multiple diverse data sources. For example, numerous systems exist for coding and annotation of video, such as ELAN [6], VCode [1], and Diver [6]. While these systems have powerful coding capabilities, they are not designed for analysis or visualization of heterogeneous data, nor do they support easily extensible automatic analysis facilities. Other systems, such as The Observer XT [7], readily support multiple types of data, but are commercial products with limited open extensibility to visualize new types of data.

FEATURES
The primary goal of DataPrism is to support interactive visualization and analysis of multiple simultaneous data streams from the same activity. DataPrism supports data import from multiple sources, including video files, audio files, log files (encoded in a variety of formats), transcript data, and paper-based digital notes.

MB'10 August 24-27, 2010, Eindhoven, Netherlands
©ACM 2010 ISBN: 978-1-60558-926-8/10/08…$10.00

Visualization

The main DataPrism interface consists of a video player and one or more timelines arrayed below the video, as shown in Figure 1. A single timeline is shown by default, but multiple timelines can be added to visualize additional data, group annotations, or view different time scales. Each timeline has a visual marker (a vertical yellow bar toward the right on each of the timelines in Figure 1) to indicate the current position in the data, which is aligned with the current video location. This marker allows the researcher to relate a notable moment in one data stream to corresponding points in other data sets.

Video files can be viewed as with a typical video viewer. The viewer displays the current frame of the video as a researcher moves through the data and allows the video to be played at varying rates. In addition, frames from the video can be viewed on the timeline below the main video viewer. Frames from the video can also be associated with annotations and displayed in a pop-up window that appears when the researcher hovers the cursor over the annotations. If the video file has an audio track, a waveform of the audio can be shown on a timeline. Multiple video files can be shown in separate windows or on separate timelines.

Log file data can currently be shown in two ways: time series graphs and map depictions of geographic position. Time series graphs are displayed on timelines in the main DataPrism window, along with dynamic scales based on the current visible data. Geographic positions can be shown on a map that is retrieved from Google Maps based on latitude and longitude, as shown in Figure 2. As with timelines, the current time position in the data can be changed by clicking on the map. In this way, a notable movement in physical space can be used as a guide to explore data in the same way as a change in a time series graph.

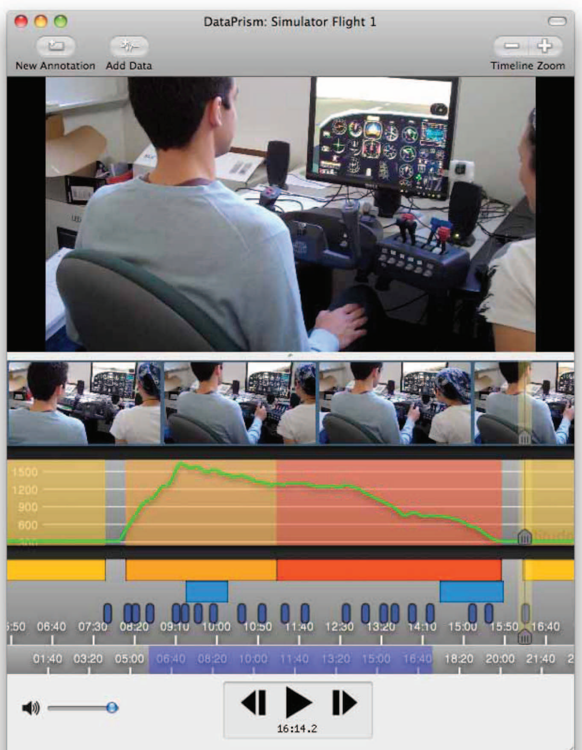

Figure 1. Main DataPrism Window, showing (from top): video, video frames, time series graph, and annotations.

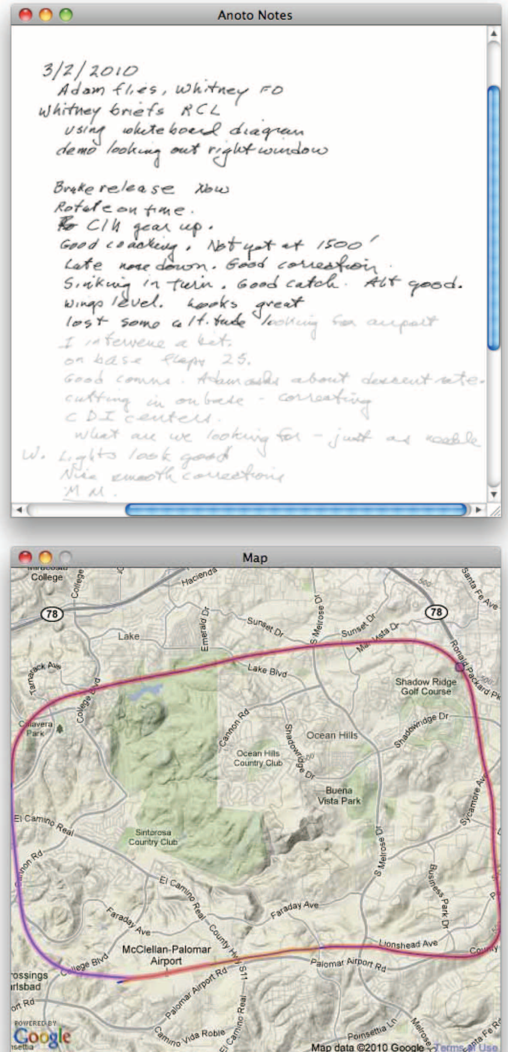

Figure 2. Digital notes (top) and map display with overlaid annotations (bottom).

Digital Pen[1] data from a researcher's paper notes can also be shown in two ways. First, markers can appear on the timeline that correspond to the time that a particular note was written. When the researcher hovers over the marker, an image of the note appears. Second, the notes as a whole are displayed, as shown in Figure 2. The pen marks are initially displayed in gray, but as the researcher moves

[1] We currently support Livescribe digital pens (www.livescribe.com) that record both pen strokes and associated audio.

through the data, the letters get darkened to indicate which notes correspond to the current time in the data. As with the timelines and the map, analysts can change the current time position in the data by clicking on particular notes to jump to the corresponding points in the rest of the data.

Annotations

Data annotations in DataPrism can be visualized on their own 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 annotations are color-coded with researcher-defined colors, and can be shown on the timelines and on maps.

Annotations can be used in multiple ways, including both continuous coding and identification of particular moments of interest in the data, as shown in Figure 3. Annotations can be simple assignments of categories or free-form text entries. Categories and possible subcategories can be defined before analysis by the researcher, and can be dynamically modified and extended during analysis. Annotations can be assigned to multiple categories and used to group, align, and filter annotations.

Synchronization

Synchronization can be a problem when separate devices are used to record multiple data sets. The ideal solution is to precisely calibrate system clocks in each of the recording devices, in which case data is automatically aligned when imported into DataPrism. Since this is not always possible with heterogeneous recording devices, interactive alignment of data is also available. With each supported data type, including multiple videos, the researcher can find a point that corresponds to another point and align the data simply by clicking on the identified point.

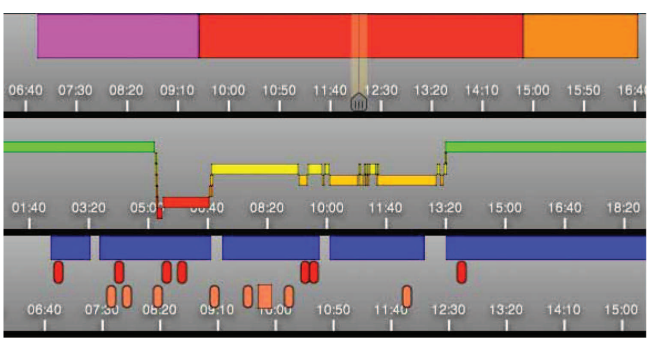

Figure 3. Different ways annotations can be coded in DataPrism. Top: Color-coded segments of data that have been manually defined to indicate phases of activity; Middle: Values with respect to a single variable assigned continuously over the course of the video, using vertical position to indicate value; Bottom: Instances of different types of activity, using vertical position to group by activity type.

Automatic Analysis

One important use of visualization is to present the results of automatic processing of data streams. DataPrism supports a data analysis plug-in system for scripts written in the Python language. These scripts can operate on any data that resides in DataPrism. The results are treated as annotations and can be visualized and filtered in the same way as normal annotations. This makes it easy, for example, for an analysis based on one set of variables to highlight a region of important activity across the entire collection of data. In the domain of flight, one type of data analysis that we have explored uses logic-based rules to identify flight segments based on well-defined flight parameters. Similarly, simple rules can be used to identify moments, such as when the airspeed is within a particular range, which might be of interest for a particular analysis.

USAGE SCENARIO

In this section, we step through a simplified description of the use of DataPrism to analyze data from a flight simulator. The scenario is based on real-world DataPrism sessions we observed. The data is from novice pilots flying the flight simulator, which created text logs of simulator variables at a rate of 5 Hz. Paper notes were made by an observer during the flight with a LiveScribe digital pen.

Analyses using DataPrism in this domain typically include the following data configuration:

- Video of pilot-simulator interaction, shown in primary video window
- Video of pilot-pilot interaction, shown in secondary video window
- Simulator log of Altitude, Airspeed, Vertical Speed, and Bank Angle variables, shown as time series
- Geographic position from simulator log, shown as a path on a map
- Notes made by the researcher, shown as annotations and as a full page

After loading the videos, simulator log, and digital notes, the researcher begins by aligning the data. When the videos were made, a clapperboard was used at the beginning to make sure that devices could be synced. The appropriate places in the movies are found, and then the data is matched to the videos by finding the point where the plane begins to move in the airspeed data and the video data. Finally, the digital notes are synced to the same point by matching a note made during the simulation at that time. This simulation run involved novice pilots, so there are many possible errors to identify in the data. On an initial visual scan of the data, the analyst notices a period of fluctuation in the vertical speed graph. By moving the current time position indicator, he gets visual reference marks of the corresponding time position in all of the other data sources. As he moves the time indicator, the position indicator on the map moves as well. This reveals that the vertical speed fluctuation occurred at the same time as a sharp turn in the flight pattern. The relevance of the fluctuations occurring during a turn is reinforced by a sudden jump in the bank angle graph at the same time. The shape of the bank angle graph reveals a sharper turn than would be needed in this situation. The combined evidence points to a lack of skill in the novice pilot for maintaining proper flight parameters during turns. By examining the visualizations from these different sources, the researcher forms a hypothesis that the cause of the poor performance might be an inability to properly manage attention between instruments depicting bank angle and vertical speed. To help evaluate this hypothesis, he looks at the video of the flight and notices the pilot's eyes seem to be locked on the artificial horizon and rarely move to the right to the location of the altitude indicator. This short scenario provides an example of how DataPrism might be used to identify an instance of poor flight performance. The important things to note are the usefulness of being able to visualize multiple data sources and how access to linked visualizations can help evaluate hypotheses that arise during analysis.

FUTURE WORK

While the current capabilities of DataPrism make it an effective tool for analysis of multimodal behavioral data, we have plans for extending it to create more powerful analysis facilities. Our design goal is to maintain the simplicity of the current interface but enable a loose coupling via a plug-in architecture with additional facilities useful for particular domains or types of analyses.

One natural direction is to explore how additional automatic analysis features might be supported. For example, computer vision techniques have advanced in capabilities and reliability to the point that they promise to be highly useful tools for aiding analysis of video data. Object recognition is one particularly promising area. It would be very useful for digital video analysis if an algorithm could automatically label all (or even most) video frames in which a particular object is present. The use of scale invariant features (e.g., SIFT and related techniques [4]) is one promising approach we are exploring. We are especially interested in supporting interactive use of automated analysis facilities. For example, since many computer vision algorithms are probabilistic, users can modify the algorithm's threshold depending on the task. For example, the threshold for object detection could be set at a low value in which virtually every frame that contains the object is detected, with the price of increased false alarms. In this case, a small amount of user intervention would be required in order to cull the false alarms from true detections. On the other hand, the object detection threshold could be set at a high value in which case there would be virtually no false alarms, with the price that in some frames the object would be present but not detected. Integrating computational analysis methods with human interaction in mind allows such decisions to rest with the analyst, and be interactively adjusted to specific analysis needs. We are also actively developing methods for integrating DataPrism with additional facilities of digital pens. A digital copy of freehand notes is a useful addition to analysis, but digital pens have several advantages over traditional pens that can be exploited during real-time observation. For example, since each stroke of the pen is time-stamped, this can eliminate the need to manually record time-stamps for later indexing into the data. In addition, there is considerable promise in linking paper notes and their digital counterparts using pen gestures [3]. We are developing a system to allow marks made in certain regions on paper (e.g., in specific areas in the margin or on customizable forms) to have special meaning and be automatically converted into annotations.

Finally, we plan to engage in further detailed studies of researcher interaction with DataPrism. While the scenario described in this paper is based on actual observations of real usage, conclusions about the efficacy of the interactive visualization techniques require additional data, including usage in other domains.

ACKNOWLEDGMENTS

NSF Grant 0729013 supported this research. We thank users of DataPrism and members of our lab for feedback and design suggestions.

REFERENCES

1. Hagedorn, J., Hailpern, J., & Karahalios, K. VCode and VData: Illustrating a new Framework for Supporting the Video Annotation Workflow. *Ext Abstracts AVI 2008*, 317-321.
2. Hodges, S., Williams, L., Berry, E., Izadi, S., Srinivasan, J., Butler, A., Smyth, G., Kapur, N., & Wood, K. SenseCam: a Retrospective Memory Aid, *Ubicomp 2006*, 177–193.
3. Liao, C., Guimbretiere, F., Kinckley, K., and Hollan, J. Papiercraft: A Gesture-Based Command System for Interactive Paper. *ACM Transactions on Computer-Human Interaction*, *14* (2008) 1-27.
4. Lowe, D. Distinctive Features from Scale Invariant Keypoints. *Intl J of Computer Vision*, *60* (2004), 91-110.
5. Pea, R., Mills, M., Rosen, J., Dauber, K., Effelsberg, W., & Hoffert, E. The Diver Project: Interactive Digital Video Repurposing. *IEEE MultiMedia*, *11*, 1 (2004), 54-61.
6. Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. ELAN: a Professional Framework for Multimodality Research. *Proc. LREC 2006*, 1556-1559.
7. Zimmerman, P.H., Bolhuis, J.E., Willemsen, A., Meyer, E.S., & Noldus, L.P.J.J. The Observer XT: A tool for the integration of and synchronization of multimodal signals. *Behavioral Research Methods, 41, 3* (2009), 731-735.