Enabling Visual Exploration of Long-term Physiological Data

Miriam Zisook[1] Javier Hernandez[2] Matthew S. Goodwin[3] Rosalind W. Picard[4]

Northeastern University Media Lab, Massachusetts Institute of Technology

Abstract

With the recent development of wearable and comfortable biosensors, larger datasets of physiological data are being collected in challenging real-life scenarios. In order to gain insight from these datasets, behavioral scientists need tools that enable agile and efficient data exploration. In this work, we designed and implemented two visualization tools for large-scale time-based datasets, and combined the lessons learned to create a more general-purpose tool.

Keywords: long-term data, contextual information, physiological information.

Index Terms: H5.m [Information Interfaces and Presentation]: Miscellaneous.

1 Introduction

For more than 150 years, psychologists and behavioral scientists have studied human emotion and behavior in controlled short-term laboratory studies. However, with the recent improvements of wearable physiological sensors, researchers have started to monitor and study emotions and behavior in more natural and uncontrolled settings over extended periods of time (e.g., several weeks or months) [1]. While this change has enabled the study of important research questions, it has also introduced new research challenges. Researchers studying this data were relying on ELAN [2] and Chronoviz [3] to visualize and explore this data, but these tools were designed for linguistics, and were not adequate for studying long-term data. Researchers also explored the data using Fry's interactive genome viewer, which could accommodate the size of the datasets, but was also not designed for temporal data [4]. As a first step towards analyzing these types of datasets, researchers need tools that enable agile and efficient data exploration. In this work, we designed and implemented two visualization tools for large-scale, multivariate, time-based datasets, and combined the lessons learned to create a general-purpose tool for data types of these kinds.

2 Case-based Scenarios

We employed an iterative participatory design approach with affective and behavioral scientists to develop two tools to enable exploration of physiological data collected over large periods of time. In particular, these researchers wanted to explore Electrodermal Activity (EDA) data and study its changes over time. EDA is a physiological measure of sympathetic nervous system arousal. Among other things, arousal mediates attention, information processing, and emotion during daily activity. The researchers were interested in two different scenarios:

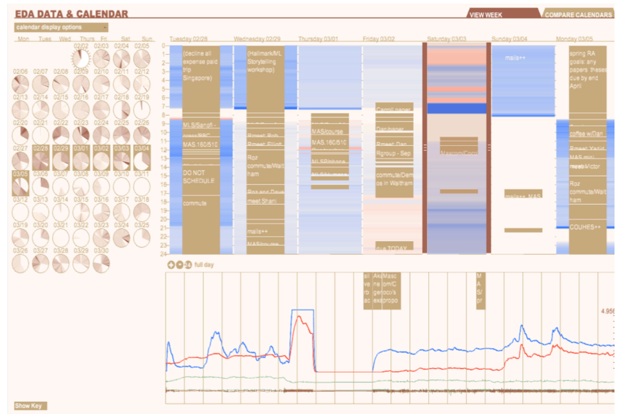

Figure 1. General-purpose tool for data exploration.

Scenario I. In the first study, researchers collected physiological data from 5 autistic children for 60 consecutive school days using the Q™ biosensor (www.qsensortech.com) that collected EDA, 3-axis acceleration, and skin temperature (all at 8 Hz) from the ankle. Furthermore, the teachers of the autistic children provided more than 6,000 contextual annotations, including settings (e.g., class), activities (e.g., reading), and behaviors (e.g., self-injury). The main need of the researchers was to easily explore associations between these settings, activities, and behaviors and physiological changes, before, during and after each event. A screenshot of our proposed solution can be seen in Figure 2 (left).

Scenario II. In the second study, researchers collected over two months of physiological data during sleep from participants using the Q™ biosensor. Study participants also completed quality of sleep journals nightly. In this case, the researchers' wanted to explore temporal relationships between self-reports and physiological changes to quickly identify repetitive patterns that may occur over time (e.g., hormonal cycles). A screenshot of the solution can be seen in Figure 2 (right).

3 Exploration of Long-term Physiological Data

After several meetings with the researchers, we identified three main challenges to be simultaneously addressed to support the aforementioned scenarios: 1) visualization of contextual data; 2) fluid exploration of multiple time scales; and 3) interpretation of physiological features. Each is discussed in turn below.

E-mail: mzisook@ccs.neu.edu[1], m.goodwin@neu.edu[3]
E-mail: {javierhr[2], picard[4]}@media.mit.edu

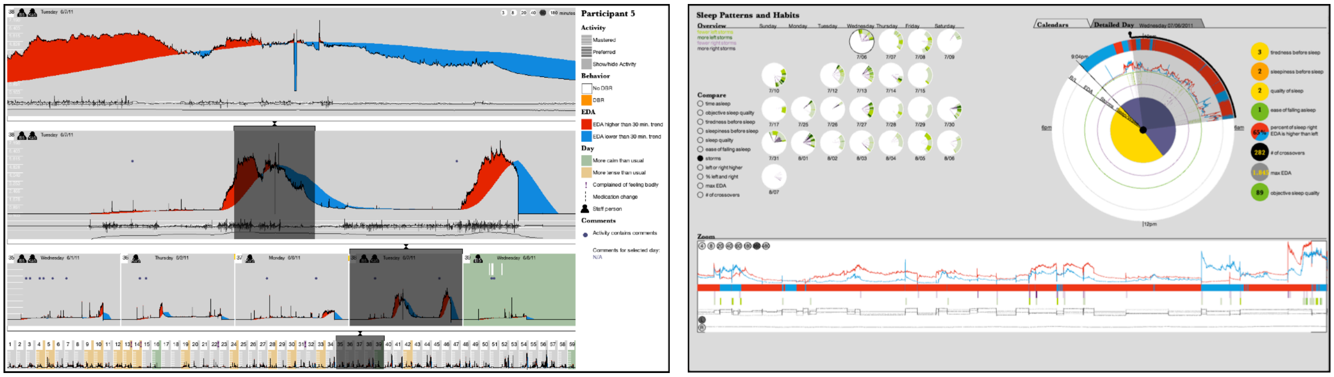

Figure 2. Screenshots of the tools developed for (left) scenario I and (right) scenario II.

3.1 Contextual Information

In order to appropriately identify the source of physiological changes, it's necessary to simultaneously visualize relevant contextual events and physiological signals. In the tool we developed for scenario I (Figure 2 left), contextual information, such as calendar events and notes about mood, is displayed by overlaying boxes time-synchronized with the EDA data. In the general-purpose tool (Figure 1) we used a similar approach by aligning physiological data with online calendar events and self-report information.

3.2 Temporal Granularity

Since physiological data may contain slow as well as fast-changing patterns, it's necessary to enable fluent exploration at different time scales. As can be seen in Figure 2 (left), four different time granularities are shown (three months, one week, one hour, and few hours). Researchers found this visualization especially useful as it allowed simultaneous exploration of different time scales. In the tool developed to address scenario II (Figure 2 right), we used small circles to represent days. While both approaches were reported to be effective, we used a mixed approach for the general-purpose tool. After several designs, we found that the time series line is most useful for viewing a single day or short time window; the vertical gradient (higher intensity indicating higher EDA amplitude) allows easy exploration of a weekly calendar; and the small circles are optimal for comparing multiple days. We also found that familiar displays of clocks and calendars helped organize time-based information and make it more familiar and intuitive. When reviewing the sleep data, a researcher was able to immediately identify when students had certain stressors such as exams, as their sleep start time drifted later and later. Prior to using this tool, it was difficult to look at enough days together to make that kind of observation.

3.3 Physiological Characteristics and Patterns

Data exploration of physiological data can be greatly improved by visualizing physiological features that help to quickly identify areas of interest. For instance, in Figure 2 (left) red and blue filled areas indicate increasing and decreasing arousal. This characteristic is especially relevant, as researchers were very interested in quickly identifying the types of physiological responses throughout daily activity. One of the researchers using the proposed tool reported to notice more periods of sharply increasing physiological arousal during certain activities that they had not been able to see before. In Figure 2 (right), the continuous bar at the bottom indicates which of the signals is higher for each time, and the information is summarized on the top-right circular graph. The use of color variation to indicate higher and lower sensor readings or features was found to be very helpful, as it allowed the quick observation of "hotspots," regions of especially high EDA, where a time series line would be harder to compare at such a high level.

4 Conclusions

In this work we created several tools to explore long-term, multivariate and time-based physiological data. The use of multiple time scales was especially helpful across tools, as well as the adjacency of annotation and biosensor data. In the future, we plan to perform a comprehensive user study to quantify what features are the most useful and easier to understand. Finding new ways to visualize physiological data and exploring how it associates with the environment, emotion, and behavior is an essential part of understanding salience for an individual person, and can enable researchers to improve the scientific understanding of psychophysiology in natural settings.

Acknowledgments

We would like to thank Akane Sano and Jean Deprey for valuable feedback on this work. This material is based upon work supported by the National Science Foundation under Grant No. NSF CCF-1029585 and the MIT Media Lab Consortium.

References

[1] Goodwin, MS, Velicer, WF, & Intille, SS (2008). Telemetric monitoring in the behavior sciences. Behavioral Research Methods, 40, 328-341.

[2] Wittenburg, P, Brugman, H, Russell, A, Klassmannm A, Sloetjes, H. "ELAN: a professional framework for multimodality research." *Proceedings of LREC*. Vol. 2006. 2006.

[3] Fouse, A, Weibel, N, Hutchins, E, Hollan, JD. "ChronoViz: A System For Supporting Navigation of Time-Coded Data." Ext. Abs. Computer Human Interaction, 299-304, 2011.

[4] Fry, B, "Computational Information Design," M.S. Thesis. MAS., MIT Media Lab, MA, 2004.