Making Sense of Wild Data: Using Visualization to Analyze In-the-Wild Video Records

Uta Hinrichs
Department of Computer Science, InnoVis Group
University of Calgary, Canada
uhinrich@ucalgary.ca

Sheelagh Carpendale
Department of Computer Science, InnoVis Group
University of Calgary, Canada
sheelagh@ucalgary.ca

ABSTRACT
In this paper we describe our use of information visualization to facilitate the analysis of in-the-wild video data. Video recording is often the method of choice when conducting in-the-wild studies. It results in highly rich and detailed data collections that can be revisited many times and analyzed from different perspectives. However, the qualitative analysis of video recordings collected in real-world settings is known as a tedious and time consuming activity, because the data can contain a large number of activity layers that have to be identified and manually extracted through video coding. We have utilized customized information visualizations to create visual representations of coded video recordings that consider particularly the temporal, social and spatial context of interactions. We describe how these visual abstractions from rich video data were valuable in various stages of our analysis process, including the cataloguing of video data, identifying research questions, in-depth analysis, and, finally, communicating our study results. We also point out various challenges that we identified in this process.

Author Keywords
qualitative video analysis, information visualization, in-the-wild studies.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

General Terms
Human Factors.

INTRODUCTION

Video recording is a popular data collection method for in-the-wild studies. Video cameras are nowadays small and unobtrusive and can capture visible and audible activities and events in great detail. Video records of activities can be revisited as often as desired and provide insights on a large variety of aspects ranging from questions around interaction with (technology) artefacts to social behaviour [1, 9]. However, analyzing video recordings from real-world scenarios can be tedious and time consuming since the data can contain dense layers of information but is only organized by time. General themes, social or spatial aspects of interactions inherent in the data have to be extracted by hand. Oftentimes, hours of video have to be painstakingly screened and catalogued, to identify general themes and snippets of particular interest. This tedious process often results in large collections of codes which can make it difficult to gain an overview of, or see relations within the data. In this paper we discuss our approach of using information visualization to abstract certain aspects from in-the-wild video data to facilitate the analysis process.

One could argue that high-cost data analysis is a general characteristic of qualitative research methods as they generate more or less unstructured data that has to be coded by hand. However, qualitative laboratory studies are usually designed around certain activities, possibly with a fixed number of participants who will work under given time constraints. In real-world settings people engage in a variety of activities, depending on their interest, background, and age [2, 4, 5]. Interactions are often intermittent and interaction times can vary greatly; people start and abandon activities as they wish [6]. Group constellations can involve both people who know each other as well as strangers and are constantly in flux [10, 12]. It is this uncontrolled nature of "in-the-wild" settings that results in rich but highly complex video data.

We turned to information visualization to facilitate our process of analyzing in-the-wild video data that was collected during a field study at the Vancouver Aquarium to investigate how visitors interacted around large direct-touch tabletop displays. We built customized visualizations that, as abstractions from the actual video data, can present overviews of patterns within particular activities of interest. In contrast to visualization tools that are available in commercial analysis software (e.g. NVIVO), our visualizations particularly consider the temporal, social and/or spatial context of interactions. This process still requires manual cataloguing and coding of the video data. However, we argue that the data generated from these coding passes can be used as basis for visualizations, which, in turn, can help to make more informed decisions on how to proceed with the analysis. We utilized our visualizations (1) to identify inconsistencies in our initial video catalogue, (2) to identify interesting video snippets for further analysis, (3) to facilitate the generation of results, and (4) to communicate and present findings. While the visualizations we present in this paper are preliminary and highly specialized for our particular study scenario and interests, we hope to open up a discussion of how visualizations can facilitate the analysis of video data collected in-the-wild.

VISUALIZING QUALITATIVE DATA

Visualizations have previously been used to help analyze qualitative data [3, 7, 8, 13, 14]. As mentioned above, some commercial tools for qualitative data analysis offer standard visualization techniques to represent the coded data (see, e.g.,

DIS'12 "Research in the Wild" workshop, June 11, 2012, Newcastle, UK.

NVIVO[1], Atlas.ti[2] or ChronoViz[3]). They mostly provide aggregated overviews of the data (e.g. in form of bar charts) but do not consider its temporal, social or spatial context.

Other approaches to visualizing qualitative data have focused on tabletop workspace activities [3, 13, 14]. These examples show the use of visualization to abstract rich interactions with direct-touch displays and, therefore, bring spatial usage patterns to the fore. Visualization has also been applied to analyze temporal sequences of visual analytics processes based on manual video coding [8]. Isenberg and Fisher have presented the Pairgrams visualization, that, based on data logs and manual video coding, highlights relations between search and reading activities within collaborative analytic reasoning processes [7]. Similar to us, they utilized visualizations in different stages of their qualitative data analysis process.

Building upon these approaches, we argue that visualizations can be particularly useful to facilitating the qualitative analysis of video recordings collected in-the-wild; in uncontrolled, real-world scenarios. In-the-wild video data differs from video recordings gathered in controlled laboratory settings in its richness and complexity. Real-world activities evolve in a rather unstructured way, influenced by the spatial and social context they occur in. This results in different, related and unrelated, layers of activities and events happening at the same time. Interactions may get interrupted and involve different variations of individual and collaborative activities. With these rich, complex and entangled activities on tape, one of the first questions is often "Where to begin?" [1, 9]. Our process aims at creating visual abstractions from the rich and complex video data to facilitate the analysis process. As part of this, we focus on representing the coded video data in context, considering its temporal, social and/or spatial relations that may be difficult to decipher from all the activities visible within the video stream. In the following we describe our approach based on the example of a field study that we conducted at the Vancouver Aquarium.

VANCOUVER AQUARIUM CASE STUDY

As part of their renovations of the *Canada's Arctic* exhibit in 2009, the Vancouver Aquarium installed two digital tabletop displays that provide information about the Arctic from an ecological, social, and political perspective. Both tables and their applications were designed by the exhibition company Ideum[4]. The Collection Viewer table (see Fig. 1, left) features photographs and videos about the Arctic as a biotope and habitat. These media items can be moved, resized and rotated in a free-form way using common multi-touch gestures [2]. The Arctic Choices table (see Fig. 1, right) features a map of the Canadian Arctic. Sliders and click wheels on the short edges of the table can be used to augment the map with visual layers that, for instance, show the sea ice change in the Arctic across the year, migration routes of different animals, political boundaries, or shipping routes.

We conducted a study at the Vancouver Aquarium two months after the initial deployment of the digital tables. The study

[1]http://www.qsrinternational.com/
[2]http://www.atlasti.com/
[3]http://chronoviz.com/
[4]http://www.ideum.com

Figure 1. Tabletop exhibits at the Vancouver Aquarium.

Figure 2. Recording visitor interaction from different perspectives.

took place on one weekend before and six consecutive days during the Christmas school holidays. The goal of the study was to investigate how visitors would interact with and around the digital tables and how they would explore content individually and/or in groups. Furthermore, we were interested in how aquarium visitors would spontaneously apply multi-touch gestures while interacting with the walk-up-and-use tables. We were not involved in the design or development of the tabletop interfaces but came in as third-party researchers to conduct the study at the Arctic exhibit. To be able to analyze visitors' activities around the tabletop exhibits in-depth, we installed two video cameras close to each table, capturing visitors' interactions from different perspectives (see Fig. 2). In particular the top-down camera view was crucial because we did not have access to the tabletop application software and, therefore, were not able to automatically log visitors' interactions on the tabletop surfaces. We recorded visitors activities for three to four hours each day, resulting in 20:38 hours of video data from each camera and 1750 visitors interacting with the two digital tables on record.

VISUALIZATIONS IN THE VIDEO ANALYSIS PROCESS

Facing this vast amount of rich video data was overwhelming. Our analysis process was largely based on the approach suggested by Heath et al. [1]: we first catalogued and reviewed all the video data, followed by a more in-depth analysis of selected video sequences. During this process, we designed custom-built visualizations to facilitate the different stages of our analysis. We describe these stages and the visualizations we built during our process in the following.

Visualizing the First-Pass Video Catalogue

Following Heath et al.'s suggestion to cataloguing the data as a first step in qualitative video analysis [1], we first coded start and end times of every interaction instance we observed in our video recordings (one instance defined by a visitor approaching the table, interacting, and leaving the table again). In this process we also added comments to interaction instances that caught our interest for future, more in-depth analysis. This first analysis pass enabled us to quantify the number and visitor types (children or adults), their interaction times, as well

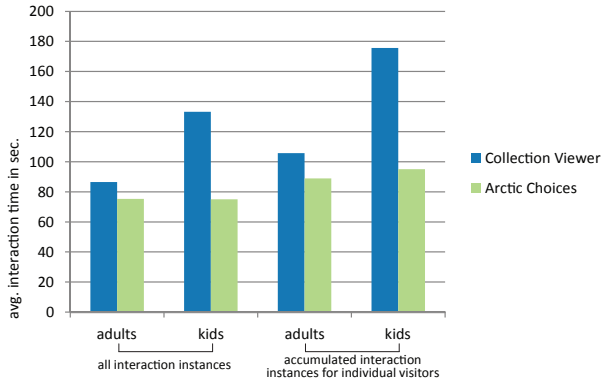

Figure 3. Average interaction times of adult and children visitors.

as repeated interactions. Visualizing these numbers in aggregation (see Fig. 3), provided some high-level insights about visitors' interactions with the two digital tables, but it did not show the interaction instances in their actual context.

We therefore designed a visualization that we call InteractionArcs. InteractionArcs show all interaction instances in sequence, also considering the different visitor types and repeated interactions (see Fig. 4 and 5). Interaction instances are represented as filled arcs, arranged on the lower part of a horizontal timeline in the order of their occurrence. Each filled arc represents the start and end of an interaction instance, with its radius representing the overall length of the instance. The colours represent visitor types: adults are shown in orange and children in blue. Arc lines above the timeline connect interaction instances by the same visitor, indicating repeated interactions. The filled black circles indicate periods where no interactions took place around the table. The labels underneath each arc correspond to codes we assigned to each visitor. While the InteractionArcs visualization may seem unconventional and highly customized toward our particular study scenario, it proved to be helpful in several ways.

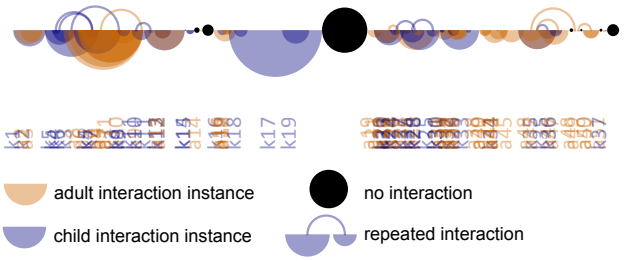

Figure 4. InteractionArcs: interaction instances in their temporal order.

Verifying the Hand-Coded Data Catalogue

The InteractionArcs visualization helped us to verify our manually created catalogue of the video data. Since we visualized all interaction instances in sequence, including "no interaction", they had to align continuously on the timeline with no gaps in-between. Gaps in the timeline could only result from coding mistakes in the video catalogue, and through the visualization we were able to easily find and correct these. Similarly, unusually large arcs caused us to double-check with the video to verify our coding.

Identifying Video Sequences for Further Analysis

The InteractionArcs visualization also revealed general patterns within the video data which was valuable for helping us select video sequences for more in-depth analysis. For instance, long interaction sequences of individuals can be easily identified in the visualization, to further explore their characteristics in the videos. The visualization also helped us to identify interaction phases where multiple visitors interacted with the tables simultaneously: overlapping of the different, slightly translucent filled arcs results in more opaque colours or in blue/orange colour mixtures, if adults and children interacted at the same time. Again, this facilitated the selection of potentially interesting video sequences for further analysis. Finally, the arc lines above the timeline revealed interesting clusters of repeated interactions (see Fig. 5, middle) which we explored in more detail.

Presenting Overviews of the Entire Data Set

The InteractionArcs visualization allowed us to present overviews of our entire data collection in a compact way. For instance, Figure 5 shows all interaction instances in sequence, as they occurred around the Arctic Choices table on a single study day. Usually, sharing data from in-the-wild studies does not go beyond transcription snippets, photographs, video stills or aggregate overviews of the data in form of bar charts (e.g. Fig. 3). Sharing raw study data is often not an option due to ethical constraints. We believe that visualizations that go beyond aggregations but consider data instances in temporal, social and/or spatial context, present an interesting and valuable alternative to sharing overviews of the entire data collection with research colleagues or clients to present results and spark discussions.

Visualizations for In-Depth Analysis

During the first high-level analysis of our video data, we noticed that there was a constant coming and going of visitors around the digital tables. Since both tables were free-standing within the Arctic exhibit, visitors approached them from all four sides, sometimes just watching other people for a while before they decided to interact themselves. During their interaction, visitors frequently shifted their positions around the table. Similarly to Marshall et al. [10] we also observed that visitors from the same group often started their interaction with the tables independently from each other, alternating between individual activities in parallel and collaborative information exploration. To analyze the positioning and movement of visitors around the table, as well as their body orientation toward the table and in relation to other people, we developed a visualization that specifically focuses on the spatial context of interactions around the table (see Fig. 6).

The visualization shows an ellipse for every visitor interacting around the table with their relative position and body orientation, including arm position. The size of an ellipse indicates visitor age (adult, child, or toddler). Ellipses of the same colour represent visitors of the same group, i.e., visitors that knew each other. For every motion that occurs around the table, a new visualization frame is generated, resulting in a sequence of position captures around the table (see Fig. 6). A timestamp in the lower left corner indicates the time the motion occurred. The bottom of each frame allows for comments (see Fig. 6), putting the movements into context.

We argue that it is particularly this *abstracted* visual representation of visitor movements around the table that is highly

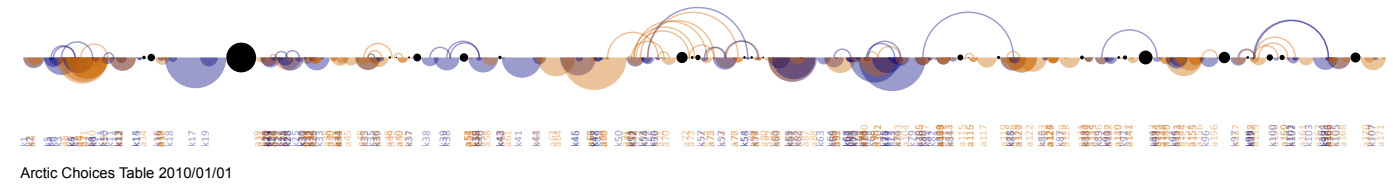

Figure 5. Overview of all interaction instances of an entire study day.

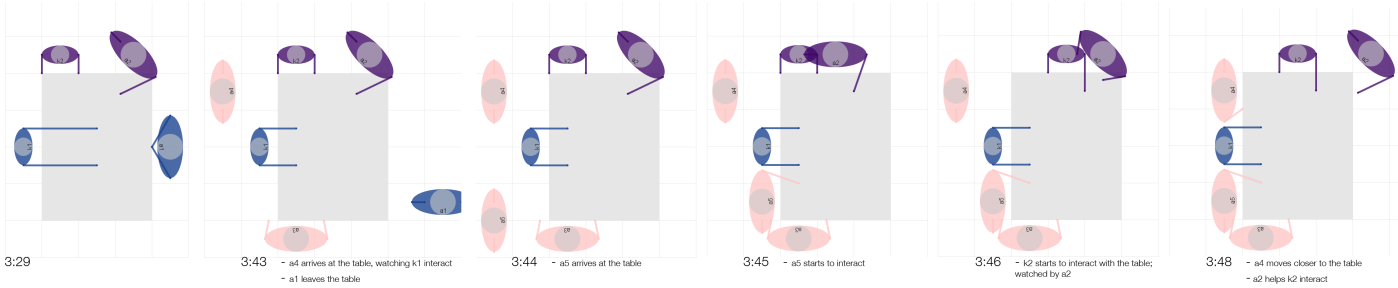

Figure 6. Position analysis of visitor movments around the table.

valuable because it allows focusing on one particular interaction aspect and how it unfolds over time. This can be useful for in-depth analysis of F-formations, for instance [11]. As mentioned earlier, the video shows ALL details of interaction in an unfiltered way. This can make it difficult to focus in-depth on particular aspects of interest. The abstraction from the video data through visualization allowed us to gain an overview of just the interaction constellations within and in-between different groups over time.

DISCUSSION

Through our explorations of utilizing information visualization to facilitate our analysis process of in-the-wild video data we have identified some challenges regarding the feasibility and value of this approach.

Visualizing In-the-Wild Video Data: Is it Worth the Effort?

As mentioned earlier, coding in-the-wild video data can be extremely time-consuming and tedious. Visualizing the coded data does not lower the time and effort required for the video analysis; it may even increase the workload considering the design of customized visualizations required to accommodate for specific scenarios and research questions. It can be uncertain how valuable the visualization ultimately will be for the analysis; several explorations into how to visualize the data may be necessary. In fact, extracting data from our videos to create the second visualization described in this paper (see Fig. 6) was highly tedious, and we are still not sure if this effort will pay off at the end. However, we believe that integrating information visualization into the process of in-the-wild video analysis will be valuable because it can help getting to know the data from different perspectives. Different perspectives can be explored on a higher level until an aspect is found to analyze in more detail. When creating the second visualization, for instance, we only coded a small amount of video data first and visualized it right away to see if it would take us into an interesting direction.

Customized Visualizations vs. Generic Visualization Tools

Related to the question of effort vs. value is the discussion of the possibility to develop generic visualization tools to support in-the-wild video analysis that would be applicable to a larger number of real-world scenarios. So far, most qualitative researchers have built their own visualizations to accommodate their particular study scenarios and research questions. One way to start thinking about more generally applicable visualization tools would be to gather common aspects of existing customized visualizations of qualitative (in-the-wild) data. However, we are not even sure if such general visualization tools would even be desirable. A quick and easy way to visualize certain aspects of video data may seem tempting but, as mentioned above, the process of designing a custom visualization itself can be considered an important part of the analysis process and may help to gain a better understanding of the data and research questions.

Automatic Logging of Visitor Data

Another approach to facilitating the visualization process of in-the-wild video recordings would be to extend methods of automatically logging activities, beyond interactions with display surfaces. Lower-cost tracking systems have become available; buttons with reflective markers could be handed out to people, for instance, as they enter an exhibition, to track their movements in the space. From an ethical point of view this approach, of course, raises a lot of questions.

CONCLUSION

Our experiences with utilizing customized information visualization to facilitate the qualitative analysis of temporal, social and spatial aspects of in-the-wild video data suggest that they are valuable in different stages of the analysis process, including the verification of codes, refining research questions, analyzing results in-depth, and, finally, communicating results to discuss with colleagues or clients. The value of visualization in all these stages is that they provide visual abstractions of particular aspects inherent in the rich but complex video data corpus. Such abstract representations can point to particular video sequences of interest and facilitate the analysis in general by providing compact but still contextual overviews of the data. Our explorations also revealed some challenges to this approach that we would like to discuss during the workshop.

ACKNOWLEDGMENTS

We would like to thank the Vancouver Aquarium, in particular Jeff Heywood, for supporting our study at the Canadas Arctic exhibit. We also thank our study participants. Special thanks go to Tony Tang and Marian Dörk for their helpful comments on this paper. We thank our iLab colleagues and our funding agencies (AITF, iCore, CFI, NSERC, and SMART Technologies) for their support.

REFERENCES

1. C. Heath, J. Hindmarsh, and P. Luff. *Video in Qualitative Research*. Sage, 2010.
2. U. Hinrichs and S. Carpendale. Gestures in the wild: Studying multi-touch gesture sequences on interactive tabletop exhibits. In *Proc. of CHI'11*, pages 3023–3032, 2011.
3. U. Hinrichs, S. Carpendale, and S. D. Scott. Evaluating the effects of fluid interface components on tabletop collaboration. In *Proc. of AVI'06*, pages 27–34, 2006.
4. U. Hinrichs, H. Schmidt, and S. Carpendale. Emdialog: Bringing information visualization into the museum. *TVCG*, 14(6):1181–1188, 2008.
5. E. Hornecker. "I don't understand it but it is cool": Visitor interactions with a multi-touch table in a museum. In *Proc. of Tabletop*, pages 121–128. IEEE, 2008.
6. E. Hornecker and E. Nicol. Towards the wild: Evaluating museum installations in semi-realistic situations. In *Proc. of Rethinking Technology in Museums*, pages 49–60, 2011.
7. P. Isenberg and D. Fisher. Pairgrams: Understanding collaborative analysis behavior with visualization. In *CHI Workshop on Analytic Provenance: Process+Interaction+Insight*, 2011.
8. P. Isenberg, A. Tang, and S. Carpendale. An exploratory study of visual information analysis. In *Proc. of CHI'08*, pages 1217–1226, 2008.
9. B. Jordan and A. Henderson. Interaction analysis: Foundations and practice. *The Journal of the Learning Sciences*, 4(1):39–103, 1995.
10. P. Marshall, R. Morris, Y. Rogers, S. Kreitmayer, and M. Davies. Re-thinking 'multi-user': An in-the-wild study of how groups approach a walk-up-and-use tabletop interface. In *Proc. of CHI*, pages 3033–3042, 2011.
11. P. Marshall, Y. Rogers, and N. Pantidi. Using f-formations to analyse spatial patterns of interaction in physical environments. In *Proc. of CSCW*, 2011.
12. P. Peltonen, E. Kurvinen, A. Salovaara, G. Jacucci, T. Ilmonen, J. Evans, A. Oulasvirta, and P. Saarikko. "It's mine, don't touch!": Interactions at a large multi-touch display in a city centre. In *Proc of CHI*, 2008.
13. S. Scott, M. Carpendale, and K. Inkpen. Territoriality in collaborative tabletop workspaces. In *Proc. of CSCW'04*, pages 294–303, 2004.
14. A. Tang, M. Pahud, S. Carpendale, and B. Buxton. VisTACO: Visualizing tabletop collaboration. In *Proc. of ITS'10*, 2010.