

Value of Using Multimodal Data in HCI Methodologies

Jhilmil Jain

Hewlett-Packard Laboratories
1501 Page Mill Road, MS#1203
Palo Alto, CA 94304 USA
jhilmil.jain@hp.com

Abstract. In this paper, I will discuss two HCI methodologies, diary study and affinity diagramming, that were carried out using multimodal data such as images, audio, video, annotated media along with the traditional use of text. I will discuss a software solution that was developed at HP Labs to conduct a multimodal diary study using three touch points: PCs, mobile devices and any type of landline/cellular phone. This will be followed by a discussion on how Microsoft StickySorter software was used to conduct multimodal affinity diagramming exercises.

Keywords: multimodal diary study, multimodal affinity diagrams, audio, video, images, text, free-hand notes, annotated media.

1 Introduction

Diary study is a methodology that asks participants to keep a diary, or journal, of their interactions with a computer system (e.g. interactions with a digital reading device) or a process (e.g. nurses record entries to uncover inefficiencies in their processes). Participants are typically required to record in real time the date and time of an event, where they were, any significant events or problems during their use of a system, and their reactions or thoughts [1]. The diary study methodology is useful when you cannot directly observe users or when you want to observe behavior over an extended period of time. It also allows users to track their own progress without a researcher standing over their shoulder. But one of the main problems with diary studies is that participants often forget to record entries or are selective in reporting.

Text based diaries do not allow for dynamic self documentation or naturalistic capturing in mobile situation. Researchers have acknowledged this problem and have created voice-mail based diaries [2], photo diaries [3], combination of photo and audio diaries [4], and combination of texting (SMS), photo and audio diaries [5]. It was found that while these may be suitable for a certain type of user study, these approaches still lack a variety of modes of expression and touch points to capture diary entries. Also, once diaries are captured, suitable means of organizing and viewing of diaries by participants or HCI researchers, and subsequently tools for aiding the analysis of diaries are still missing or are not well integrated.

Researchers conducting studies to observe mobile behavior [6] or group behavior [7] have respectively observed the need for a methodology that allows flexibility of

expression from one or more devices, and one that allows seamless collaborative creation of the diary by multiple participants.

To address these issues, I propose using multimodal diaries that allow entries to be made in multiple formats and from various touch points. I hypothesize that these will be ideal for the studying user interactions and behaviors in the following scenarios:

- Multiple locations – e.g. shopping lists created at home, at work, when the user is out and about
- Various touch points – e.g. accessing bank accounts or information on the Web, ATM, or using mobile device
- Collaboration – e.g. creating a travel plan with friends and family
- Time line specific – e.g. using Web or mobile devices for trouble shooting activities
- Complex processes – e.g. planning, organizing, scheduling and orchestrating an event such as corporate events or wedding planning
- Iterative processes – e.g. using Web or mobile devices for real time visual analysis of data
- Online research/web activity
- Mobile activity

In section 2, I will describe a multimodal diary solution that was developed at HP Labs. Please note that at the time of writing this paper, this solution was not yet used by researchers to conduct a formal user study.

Affinity diagramming is a methodology that is used to create hierarchy of observations such that they can be grouped into trends. This enables us to uncover tools, terminology, methods, strategies used by the participants [1]. Traditionally, text-based notes are used for conducting affinity diagramming. For analyzing multimodal diaries, I have extended the affinity diagrams to contain audio, video and still images as well such that richer data sets can be easily analyzed and grouped while preserving the context of creation. In section 3, I will discuss how I used Microsoft StickySorter [8] to conduct multimodal affinity diagramming exercise.

2 Multimodal Diary Study

At HP Labs, a multimodal diary solution was implemented that enables the capturing of diary elements using different modalities or formats such as structured text, audio, still images, video, unstructured scribbles and annotated media (i.e. annotating images, video frames etc.). By using the annotation feature, users can add position-based audio, video, images or text annotation on top of an image.

Diary entries can be recorded from three primary sources:

- 1) On the PC, by using a web browser toolbar (fig. 1 and 2),
- 2) On any phone (landline or cell phone), by dialing into a number and leaving an audio message (when users do not have access to the Web on the phone), or
- 3) On any mobile device, by downloading a mobile client on their PDAs and accessing the service using the data network provided by the cellular carrier or any wireless network (fig. 3 and 5) or by using the MMS (Multimedia Messaging Service) service.

These multimodal diaries can be accessed, edited and organized via the Web on any PC (fig. 6) or mobile phone (fig. 4).

2.1 Capturing Diary Entries

Participants need to install a Web browser toolbar (fig. 1) to record entries from their PCs. During this process, the participant is required to choose a username and password to use the service from a Web browser, and a phone number to use the service from a landline or cell phone.

To record an entry from the Web, the participant needs to simply highlight the information, (this could be in the form of text, image, audio or video clip) on the Web page (fig. 2) and click on the relevant diary name (e.g. “BBQ Grill”). Behind the scenes, a request is sent to the Web server with the Web page URL and the date and time of the capture. There are no restrictions on the type of information (i.e. text/image/audio/video) or the website – thus any information from any website can be added to the diary.

Audio entries from a phone can be added to the diary by calling a predefined phone-number to record a voice message. In the backend, the solution extracts keywords from the audio and adds it to the original recorded audio, along with the date and time information of the recording.



Fig. 1. Closeup of the browser toolbar that allows users to capture diary entries from the Web

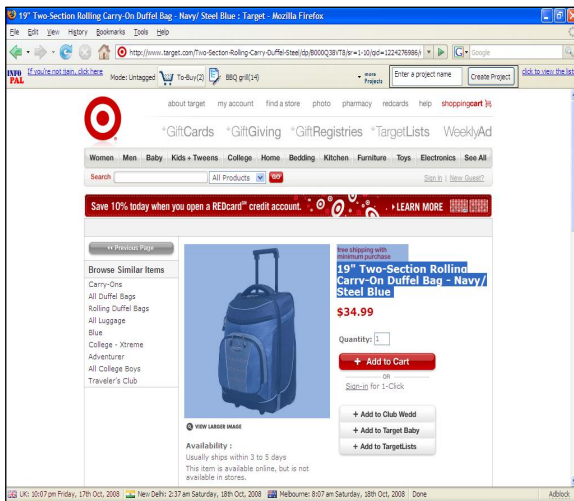


Fig. 2. To capture information (such as text and images) from the web, highlight relevant information and click on the diary name to add it. In this figure, the image and the text on the right to it are being highlighted such that they can be added to a diary.

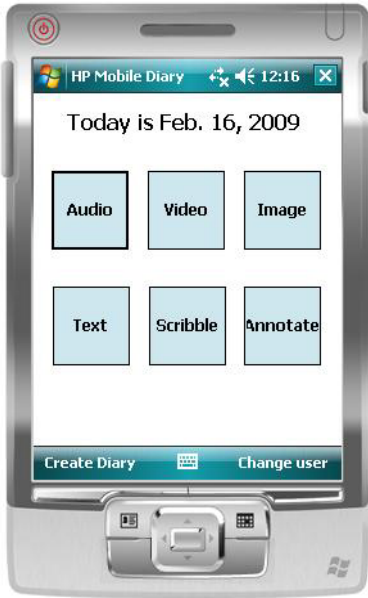


Fig. 3. Mobile client for capturing diary entries from handhelds



Fig. 4. Diary entries can be viewed on mobile devices as well



Fig. 5. Example of multimodal annotation tags: positional (video, audio and structured text) and non-positional sketching added on a still image by a user



Fig. 6. Multimodal diaries can be accessed and organized via a Web based service. Samples diaries in the figure contain audio, images and text based.

Using the mobile client, participants can additionally annotate still images or video frames by using multimodal tags (fig. 5). These tags consist of: structured text (entered via the mobile device keyboard), free-hand notes (written on the screen using the stylus/finger), audio, and video. Each of these types of tags are position based, that is if added on coordinate $x=10$, and $y=25$, their position is retained. A non-positional tag, such as sketching on the image itself is also available. Using both positional and non-positional tags, participants have the capability to easily capture rich contextual data.

The participants can access the aggregated multimodal diary from any Web enabled device – on the PC (fig. 6) or on the mobile device (fig 4). Participants can even collaboratively add to the same diary by using the same login information (from the PC) or mobile number (from the phone). Additionally, participants can even right click on the diary and share it with one or more people.

Still images, audio and video entries can be created via MMS by using preexisting software on the cell phone. For using this option the participants must have a data plan on their mobile device from their network providers. The MMS message sent from the phone is parsed by out backend solution such that the multimedia file, date, time (and optionally text) information is extracted and added to the appropriate diary. The requirement is that the diary name is the subject of the MMS message.

Alternatively, participants can download and install a mobile client to add multimodal information to the diary (fig 3). To use the mobile client, participants can either use their network provider's data plan or any wireless connection or Bluetooth connection. The HP multimodal diary solution's mobile client is network agnostic.

2.2 Managing and Organizing Diary Entries

Entries captured from the Web or the phone can be tagged. Each entry can have one or more tags such that notes can be sectioned into categories for easy organization. These tags can be created by the participant of the study to categorize the entries. Purely text based entries (like in a traditional diary study) can also be made from the Web or a mobile phone.

For all entries made from the Web, the date, time and URL are automatically captured. Similarly, all data captured from the mobile phone is automatically tagged with date and time stamp, location stamp (using GPS information), and all voice mails have a date and time stamp as well. Participants of the diary study no longer have to worry about recording the date and time of entries.

All diary entries can be easily sorted based on time, input format (e.g. audio, video, images etc.) and tags.

Each entry is treated as a block in the diary and can be moved from one diary to another or re-ordered within each diary using a simple drag-and-drop operation. Double clicking on the note allows the user to edit them. Notes can be deleted and cloned as well.


The solution also allows users (participants of the study or the researchers) to select one or more entries of the diary for publishing them in various formats (e.g. pdf, word). Any multimodal entry type (e.g. images, video, audio, text) can be included in the published document (fig. 7). For audio entries, keywords from the file are extracted and displayed on the document. For video entries, key-frames from the file are extracted and displayed on the document. There are a number of parameters available for customizing the layout of the document: number of columns, include data source of the entry, organize entries by time or tags etc.

2.3 Implementation Details


The Web toolbar extension was implemented for the Firefox browser. Regular POTS (plain old telephone service) or Cellular carrier was used for connection between the customer and the system, over any phone line. An audio-blogging website was used to capture audio lists and Dragon Naturally Speaking was used for speech to text conversion. Any MMS enabled cell phone can be used to send images to the system. SQL Server 2005 was used as the database to store intent. Microsoft IIS was used as the web server to host shopping list interfaces on the Web. Apache Tomcat was used as the web server to host the sign-up interface, sign-in interface, toolbar extension, diary interface on the PC and phone, and all updates made to the diary database. All software was implemented using Microsoft .NET 2.0 framework.

Diary: BBQ Grill
Jhilmil Jain

Printed on February 20, 2009



• CharmGlow 4 burner



• Ducane 3 burner -1 side



• One chef listened quietly while I ranted about the virtues of gas grilling, waited until I was finished, and then said, "Yes, charcoal grills are more work and what type of grill you use will depend on your lifestyle and the amount of time you have available, but in the end you can't compare gas grilling to charcoal grilling. Forget about it, charcoal wins hands down."

http://www.elucotgourmet.com/barbeque_grills.htm

Gas grill

- They provide strong, even heat, are far easier to light (just turn on), and offer excellent temperature control.
- Look for a gas grill that's sturdy and heats quickly. Porcelain-coated grates help prevent sticking, and extra features, such as side burners or rotisseries, are interesting options.

http://www.foodnetwork.com/food/et_al_grilling/article/0,10772,FOOD_9872_1772856_00.html


Charcoal grill

- To light a charcoal grill, skip the lighter fluid, which can give food a strange chemical aftertaste.
- Get a chimney—a cylindrical, metal apparatus that holds briquettes in a light pile so they heat quickly. After heating, pour the coals into your grill and get cooking.


http://www.foodnetwork.com/food/et_al_grilling/article/0,10772,FOOD_9872_1772856_00.html

Gas grill

- Figure 1: Gas grills are easy to clean. (Photo courtesy of Brilmaster.)



<http://www.dummies.com/WileyCDA/DummiesArticle/0,4415.html>

powered by 

Page 1

Fig. 7. One or more multimodal items from the diary can be chosen by the participant or the HCI researcher to be published in various formats such as pdf. Various templates for formatting are also available.

3 Multimodal Affinity Diagramming

Using this solution, researchers can organize and categorize the diaries as they see fit and subsequently they can automatically generate .CSV (comma separated values) files that can be opened and modified using Microsoft Excel. The .CSV file is automatically generated to contain the following entries since my goal was to conduct the affinity diagramming exercise using Microsoft StickySorter:

- *Text (Field 1)* – text entered by the diary participant. For audio entries, this field contains the keywords extracted from the file. For the rest of the formats, this field contains any textual annotations added by the participant.
- *Hyperlink* – If the note is an audio, video or an image, the source of the file is automatically added such that clicking on the note in StickySorter will open the file.

- *Tag* – the categories/tags that were created by the participants.
- *Touch point* – location where the entry was created. For Web entries, it contains the URL of the website from where the information was collected and for mobile entries, it contains the GPS location (latitude and longitude). If a database of latitude and longitude with corresponding address was created by the researcher, then this field contains the address details.
- *Creator* – name or identification of the participant that create the entry
- *Date Created* – date of diary entry
- *@X Position and @Y Position* – these will be filled in by the Microsoft Sticky-Sorter when researchers organize the notes on the display.
- *@Note Color* - these will be filled in by the Microsoft StickySorter when researchers create note hierarchies. By default, all notes are created “Yellow”.
- *!Group* – name of the group that the note will be assigned to (e.g. in figure 8, two out of the four notes have been assigned to the group “Multimodal Diary”)
- *@Note ID*- automatically filled in by Microsoft StickySorter.

Figures 8 and 9 show screen shot of an example .CSV file created by the solution.

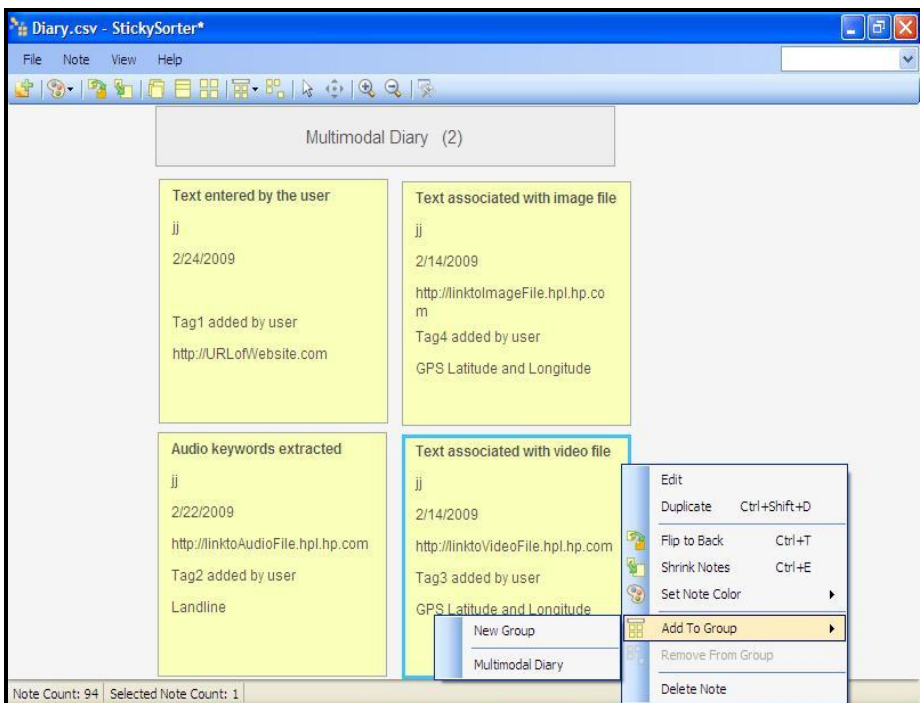


Fig. 8. Screen shot of Microsoft StickySorter displaying the .CSV file create by the solution. Four sample notes have been created, representing four entries created by a participant. Top left displays a text entry, top right displays an image entry, bottom left displays an audio entry and bottom right displays a video entry.

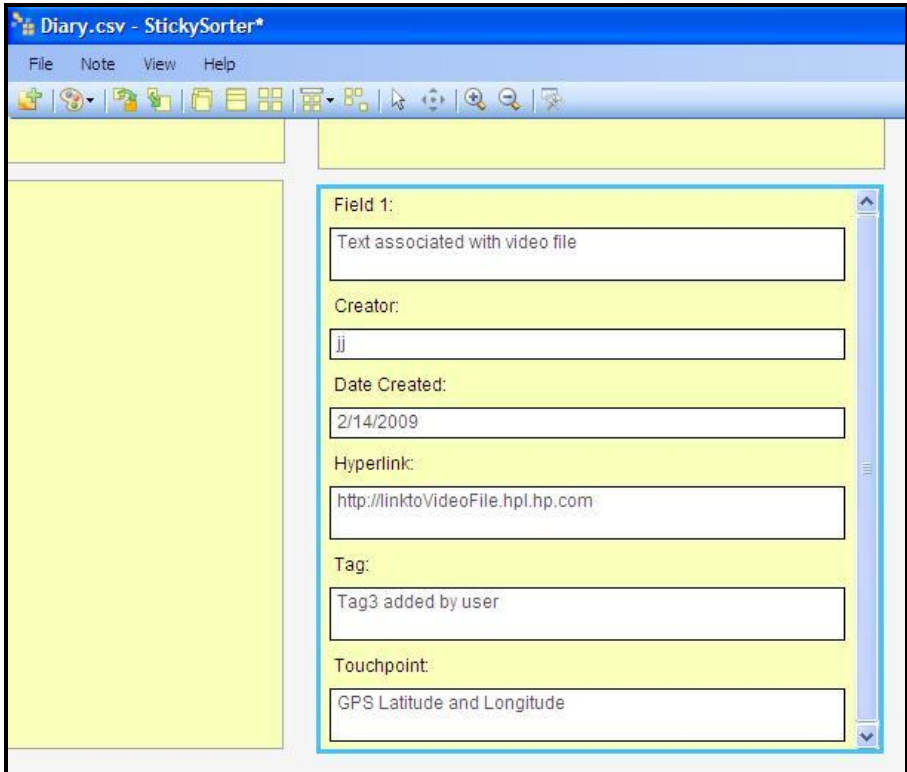


Fig. 9. Clicking on a note makes it editable. Using Microsoft StickySorter notes can be added to groups, other categories within a note can be added as well, color of notes can be changed to reflect ranking etc.

4 Conclusions

In this paper, I have described a multimodal diary solution that was developed at HP Labs. This methodology allows participants using a diary study the flexibility of expression using multiple modes of input such as text, images, audio, video, free-hand notes and annotated media. The participant is not bound to a device or a note-book to create, access or edit entries since a variety of touch points (such as PCs, mobile devices and landline) are supported by the system. Using this methodology and solution a variety of scenarios that were previously not ideal to be studied using diary studies can now be considered by HCI researchers.

Multimodal data gathered by the diary can be seamlessly converted into a .CSV file that can be imported into Microsoft StickySorter for conducting a multimodal affinity diagramming exercise.

The next steps of this project are to recruit HCI researchers to conduct multimodal diary studies using this tool, such that it can be further enhanced. More research is required to develop methodologies for analyzing qualitative multimodal data that is captured over time and requires to be compared over a time period.

Acknowledgments. I would like to thank Catherine Courage for providing helpful feedback on initial drafts, Riddhiman Ghosh for developing the RSS feed for the voice mail feature of the solution, and my managers Mohamed Dekhil and Meichun Hsu for sponsoring this project.

References

1. Kuniavsky, M.: *Observing the User Experience: A Practitioner's Guide to User Research*. Morgan Kaufmann, San Francisco (2003)
2. Palen, L., Salzman, M.: Voice-mail diary studies for naturalistic data capture under mobile conditions. In: *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work. CSCW 2002*, New Orleans, Louisiana, USA, November 16 - 20, 2002, pp. 87–95. ACM, New York (2002)
3. Van House, N.A.: Interview viz: visualization-assisted photo elicitation. In: *CHI 2006 Extended Abstracts on Human Factors in Computing Systems, CHI 2006*, Montréal, Québec, Canada, April 22 - 27, 2006, pp. 1463–1468. ACM, New York (2006)
4. Carter, S., Mankoff, J.: When participants do the capturing: the role of media in diary studies. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2005*, Portland, Oregon, USA, April 02-07, 2005, pp. 899–908. ACM, New York (2005)
5. Brandt, J., Weiss, N., Klemmer, S.R.: Txt 4 l8r: lowering the burden for diary studies under mobile conditions. In: *CHI 2007 Extended Abstracts on Human Factors in Computing Systems, CHI 2007*, San Jose, CA, USA, April 28 - May 03, 2007, pp. 2303–2308. ACM, New York (2007)
6. Sohn, T., Li, K.A., Griswold, W.G., Hollan, J.D.: A diary study of mobile information needs. In: *Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems, CHI 2008*, Florence, Italy, April 05 - 10, 2008, pp. 433–442. ACM, New York (2008)
7. Hyldegård, J.: Using diaries in group based information behavior research: a methodological study. In: *Proceedings of the 1st international Conference on information interaction in Context*, Copenhagen, Denmark, October 18 - 20, 2006. IiiX, vol. 176, pp. 153–161. ACM, New York (2006)
8. Microsoft StickySorter,
<http://www.officelabs.com/projects/stickysorter/Pages/default.aspx>