Informed Browsing: Scaling Up Co-Experienced Access to Digital Media

Otmar Hilliges
Media Informatics Group
University of Munich
Amalienstr. 17, 80333 Munich, Germany
otmar.hilliges@ifi.lmu.de

ABSTRACT

Recent progress in the field of multi-touch interactive surfaces makes it imaginable that our daily environments will become interactive and acquire new functionalities provided by computing capabilities. For example, walls, doors and tables could double as displays and accept input. The spatial and social affordances of tables predestine them for sharing media artifacts with friends and family. Interactive tables are subject to different design constraints than conventional interfaces. In this paper I present individual interfaces as primary interaction style for tabletop multimedia applications. Furthermore, I present the concept of informed browsing as a mean to deal with large amounts of data in this context.

Categories and Subject Descriptors: H.5.2 [**User Interfaces**]: Graphical User Interfaces (GUI), Input devices and strategies. H.5.3 [**Group and Organization Interfaces**]: Collaborative computing

Additional Keywords and Phrases: tabletop, single display groupware, information visualization, interactive surfaces, informed browsing

INTRODUCTION

Digital media (photos, audio, video) has practically replaced its analog counterpart in almost all aspects of life. Technological advances have led to ever increasing amounts of data in both professional and private contexts, since creation, reproduction and storage costs have been significantly reduced. In response to this, a variety of software for browsing, organizing and retrieving digital media has been developed.

Specifically in regard to digital photo tools, two strategies to achieve these goals have matured over time: First, advanced grouping methods and/or zooming interfaces [16] are applied in order to maximize screen real estate and to show as many pictures as possible at one time. Second, search engines help users to retrieve specific photos in a more goal-driven way. Since photos are perceived through the content shown, effective search relies on textual annotations or tagging with automatically derived meta data [3, 20]. Tagging-based approaches are very popular for online sharing of pictures (e.g., Flickr.com, Zoomr.com). However, recent studies suggest that users are reluctant to make use of annotation techniques [15] and might not even want to perform query-based searches [11] in their own personal collections.

These approaches share one property: They have been designed and optimized for single-user interaction on standard desktop computers; sharing of pictures thus happens via e-mail or through websites and web services. Research dealing with actual photo usage finds that people strongly prefer co-located sharing of pictures with friends and family over sharing photos remotely [4, 9].

The desktop PC is not well suited for co-present collaboration since the size and orientation of standard displays impede face-to-face communication. Desktop systems do not afford mutual eye contact and body language as well as other properties important for verbal and non-verbal communication. Furthermore, PCs lack the tangibility of physical media, which is also very important for co-experiencing digital media.

TABLETOP MULTIMEDIA ACCESS

The technology for large interactive surfaces has rapidly matured in the last years. Interactive tabletop systems, in particular, have come close to the point where we can expect them to be productized and marketed, hence impacting our daily lives more significantly (e.g., SmartTech's DViT [1], DiamondTouch [2]). Further research with frustrated internal reflection and optical flow analysis has resulted in systems that drive technological advancements even further [5, 21] and open up new and interesting opportunities for tabletop applications by allowing the recognition of multiple touchpoints simultaneously. Finally, Microsoft has recently announced a multi-touch tabletop system as a commercial product [2].

Not surprisingly, the popular application field of digital photography has served as a scenario for tabletop research. Hinrichs et al. [8] have studied the effects of their "interface currents" on the collaborative use of photo collections. Morris et al. [13, 12] have explored how the orientation and distribution of control elements influence group performance on interactive tables using a photo tagging/searching scenario. An extensive body of literature about consumer behavior regarding digital (but also printed) photos has emerged in recent years [1, 4, 9, 15]. All studies confirm that users share a strong preference for browsing through their collections as

[1] http://www.smarttech.com/DViT/
[2] http://www.microsoft.com/surface

UIST '07, October 07–10, 2007, Newport, Rhode Island, USA
ACM 978-1-59593-679-2/07/0010

opposed to explicit searching. The Personal Digital Historian (PDH) [19] is a tabletop application that enables users to share pictures based on the four Ws of storytelling (i.e., Who, Where, What, When). However, to render this support possible the PDH requires an extensive set of meta data, which is seldom found in personal image collections (cf. [15]).

Recent research suggests that tabletop interfaces are subject to different requirements than desktop UIs. Foremost because orientation and spatial layout influence communication and interaction in tabletop collaborative work [10, 17, 18]. In consequence tabletop photoware has different design needs than desktop photo management tools. The peculiarities of single display groupware (SDG) make the adaptation of many conventional strategies for media tools difficult or impossible. For example, global layout strategies of photos can have negative effects on visibility and reachability for users seated on different sides of the table.

In my dissertation I am developing new interaction and visualization techniques that effectively support the browsing, organizing and co-located sharing of digital media on interactive surfaces. My current approach is twofold: 1) In order to address orientation, sharing and access allocation issues, I propose to use *individual interfaces* to grant multiple users access to shared information landscapes while maintaining the fluid social transitions between group and individual work which are essential for face-to-face collaboration. 2) To cope with large amounts of data and prevent information overflow, I want to enrich the browsing process by *informed browsing*: a combination of user interface design and information retrieval techniques that augment exploratory and browsing-based media retrieval as well as face-to-face sharing of media.

Individual Interfaces

On the desktop only one person can directly interact with a standard PC, and all other collaborators are degraded to merely follow the leaders' actions. In contrast, SDG enables users to share resources and interact in a truly parallel way. They might work concurrently or occasionally join their efforts to reach common goals. An implication of multiple users sharing displays and information is that information cannot be manipulated on a global scope by one user without potential interference with the interests of other users. Hence data manipulation should be limited to a local scope without restricting collaboration amongst users.

Figure 1: Screenshots from the Living-Room project [7] Left: Annotating pictures. The extended region can be used for (hand) writing in annotations. Right: Filtering the images of one pile. Previews of matching photos are shown on the outskirt of the semi-transparent overlay.

To address the issues outlined above, I propose the concept of individual interfaces as views onto shared information landscapes. These interfaces are portable and can be freely positioned by its user. Currently I use physical handles which, upon contact with a display surface, extend into a virtual, semi-transparent overlay (see Figure 1). Thus, the spatial limitations with which traditional menus have to cope are addressed. Since each user has a replicated interface, all the controls needed are always in place, and actions can be applied in a fluid manner. In contrast to WIMP controls, no additional movement of the hand (to and from a menu) is necessary. Also the problems of positioning, orientation and order of access can be addressed by replicating the controls instead of having a centralized menu.

Second, the concept of personalized interfaces can support collaboration by offering explicit ways of communication. The immediate hands-on nature of individual interfaces, for example, aids in non-verbal communication through natural hand gestures and body language. Users' interaction with their individual interfaces also helps reveal to one another what the task at hand is, and also helps collaborators to understand which parts of the information are presently being inspected ("look at this").

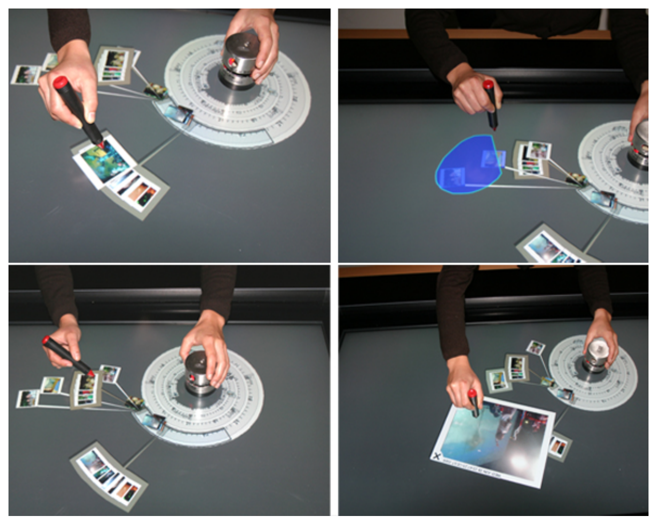

Figure 2: Co-Experienced media access. Users have replicated controls as well as re-orientable and replicated data artifacts to enable truly parallel interaction and sharing of pictures.

As an example system I have developed Photohelix (see Figure 2), a photo-browsing application tailored for filing, co-experienced browsing and sharing of pictures on an interactive digital tabletop (see also Video[3]). It is based on the individual interface concept. The system uses the notion of time and events to organize and pre-cluster collections. Events are represented as image piles on a helix-shaped calendar. The helix is generated dynamically and spans the entire collection from the oldest pictures on the inside to the newest pictures on the outside. The inner windings of the helix are rendered

[3] `www.medien.ifi.lmu.de/team/otmar.hilliges/files/ph.avi`

much smaller than the outer windings, which correlates with findings that suggest that newer pictures are accessed more frequently [9] and thus should be represented in more detail. Events and pictures are accessed, manipulated and inspected using a hybrid, bi-manual interaction technique. The non-dominant hand operates a physical handle to position and control the calendar view (rotation adjusts the current time setting). The dominant hand is used to inspect and modify events as well as individual pictures for browsing and sharing purposes. A user study with 20 participants yielded positive results and encouraging feedback. The detailed findings will be reported in a forthcoming paper.

However, Photohelix was designed and implemented in order to assess the validity of the identified requirements and design considerations – not as a system working under realistic circumstances. Hence, scalability is an issue in the current state of implementation. Photohelix does not optically condense the information shown at any given time further than pre-grouping images into piles on the helix. Therefore, the current approach suffers from visual clutter once these groups contain more than approximately 30 pictures each. This would be rather frequent under realistic circumstances, for example, many pictures taken at a wedding (i.e., over a short period in time). To address these limitations and in order to scale applications that apply the individual interfaces concept to realistic sizes I have developed the concept of *Informed Browsing*.

Informed Browsing

Throughout the body of literature, a set of typical activities performed with media collections can be found. Future applications should try to support these activities, which are: 1) *Browsing* - Users look at pictures from different time periods, possibly to revive old and forgotten memories. 2) *Selecting* - A repetitive activity in which users go through their collections and decide which items to keep and which to get rid of. 3) *Sharing* - Often the ultimate usage of media at the end of its lifecycle. This can be performed remotely via e-mail or websites but also (and preferably [1]) co-located for communication and storytelling, such as updating friends and family about recent events. 4) *Filing* - The task of sorting media into folders or albums.

I think that automatic image analysis can help support users in these tasks, especially when these technologies are carefully instrumented to support the users' semantical understanding of images, instead of stubbornly collecting as much data as possible to be used in a search-by-similarity approach – an attempt whose results might in the end be hard to understand for users [14].

Informed browsing is a combination of information retrieval technologies and interface design. It incorporates three major principles: *overview at all times*, *details on demand*, and *temporary structures*. *Overview at all times* can be achieved through automated pre-clustering and (local) screen real estate maximization. *Details on demand* requires interaction techniques that allow users to retrieve detailed information quickly. And the possibility to create *temporary structures* eases the process of comparing pictures for filing and selecting. *Temporary structures* also allow users to create arrangements for sharing and storytelling (e.g., all beach shots from the last vacation or a playlist of current favorite songs) without disrupting the long-term organization of the media collection.

	Browsing	**Selecting**	**Sharing**	**Filing**
Quality	omit "bad" pictures	sort by quality	share only HQ pictures	global pre-clustering
Color	group by color	find over- or under-exposed images		global pre-clustering
Texture	group by texture			
Edges	group by geometry	auto rotate		
Geo-metric shapes	only show certain objects		pre-cluster by object	global pre-clustering
Face detec-tion	only show portraits	drop beveled portraits	separate landscape from portrait shots	global pre-clustering

Table 1: Informed Browsing: supporting different browsing activities with information retrieval technologies

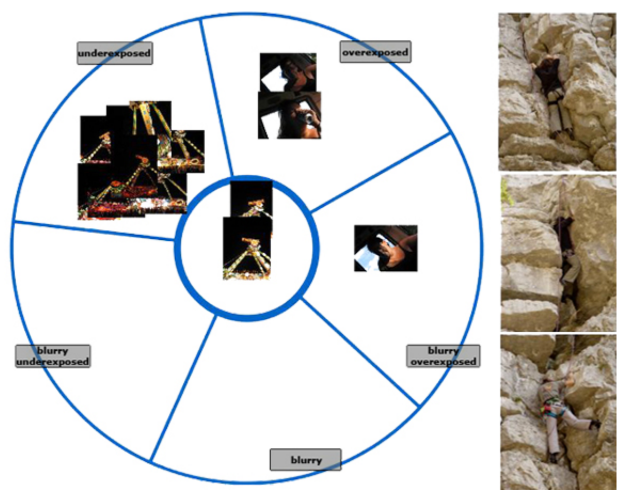

Figure 3: Quality based presentation of pictures [6]. Left: Isolating "bad" pictures. Right: Selecting the best shot out of a series of similar motives.

Up to this point I have developed a framework that extracts several attributes of digital images by analyzing their content. Table 1 summarizes the attributes that currently can be extracted and how they can be used to augment the graphical representations and hence the browsing experience of picture collections. It is important that these modifications are not applied globally (nor are they applied constantly) but instead the data is used to selectively improve the browsing experience when necessary. For example, one would only use automatically derived quality data when the task at hand is to decide which images to keep and which not (see Figure 3). Currently I am conducting several experiments in order to find out whether to apply certain functionalities in an automatically derived, context sensitive manner or by enabling users to explicitly activate and deactivate additional functionalities.

CONCLUSIONS AND FUTURE WORK

To this end I designed and implemented different systems to browse, organize and share digital photo collections [7] based on the concept of individual interfaces. Results from early user studies suggest that the presented prototypes meet the special requirements imposed by the informal and highly dynamic nature of the photo handling process. The evaluations also revealed that our design considerations have a positive impact on the usability and perception of tabletop photoware. The general feedback from users suggests that interfaces like these might help to close the gap in emotional attitude toward digital photos versus their printed counterpart.

However, the evaluations have also uncovered that advanced interface design strategies are required in order to deal with ever increasing amounts of digital media artifacts. I have proposed the notion of informed browsing as an conceptual model of how information retrieval technologies can improve the media browsing experience. Furthermore I have developed a modular framework that incorporates state-of-the-art image analysis algorithms.

In the future I plan to extend both the conceptual and the technical framework to incorporate audio and video as well. Finally I plan to conduct several user studies in order to understand more deeply how informed browsing can support the co-experienced access to multimedia collections.

REFERENCES

1. A. Crabtree, T. Rodden, and J. Mariani. Collaborating around collections: informing the continued development of photoware. In *Proceedings of CSCW '04*, pages 396–405, 2004.
2. P. Dietz and D. Leigh. Diamondtouch: a multi-user touch technology. In *Proceedings of the ACM UIST '01*, pages 219–226, 2001.
3. S.M. Drucker, C. Wong, A. Roseway, S. Glenner, and S. De Mar. Mediabrowser: reclaiming the shoebox. In *Proceedings of AVI '04*, pages 433–436, 2004.
4. D. Frohlich, A. Kuchinsky, C. Pering, A. Don, and S. Ariss. Requirements for photoware. In *Proceedings of CSCW '02*, pages 166–175, 2002.
5. J.Y. Han. Low-cost multi-touch sensing through frustrated total internal reflection. In *Proceedings of UIST '05*, pages 115–118, 2005.
6. O. Hilliges, P. Kunath, A. Pryakhin, H.P. Kriegel, and A. Butz. Browsing and Sorting Digital Pictures using Automatic Image Classification and Quality Analysis. In *Proceedings of HCI International '07*, July 2007.
7. O. Hilliges, M. Wagner, L. Terrenghi, and L. Butz. The Living-Room: Browsing, Organizing and Presenting Digital Image Collections in Interactive Environments. In *Proceedings of IE '07*, September 2007.
8. U. Hinrichs, S. Carpendale, and S.D. Scott. Evaluating the effects of fluid interface components on tabletop collaboration. In *Proceedings of the working conference on Advanced visual interfaces*, pages 27–34, 2006.
9. D. Kirk, A. Sellen, C. Rother, and K. Wood. Understanding photowork. In *Proceedings of CHI '06*, pages 761–770, 2006.
10. R. Kruger, S. Carpendale, S.D. Scott, and S. Greenberg. How people use orientation on tables: comprehension, coordination and communication. In *Proceedings of the International ACM SIGGROUP conference on Supporting group work*, pages 369–378, 2003.
11. T. Mills, D. Pye, D. Sinclair, and K. Wood. Shoebox: A digital photo management system. Technical report AT&T Laboratories Cambridge., 2000.
12. M.R. Morris, A. Paepcke, and T. Winograd. Teamsearch: Comparing techniques for co-present collaborative search of digital media. In *Proceedings of TABLETOP '06*, pages 97–104, 2006.
13. M.R. Morris, A. Paepcke, T. Winograd, and J. Stamberger. Teamtag: exploring centralized versus replicated controls for co-located tabletop groupware. In *Proceedings of CHI '06*, pages 1273–1282, 2006.
14. K. Rodden, W. Basalaj, D. Sinclair, and K.R. Wood. Does organisation by similarity assist image browsing? In *Proc. CHI '01*, pages 190–197, 2001.
15. K. Rodden and K.R. Wood. How do people manage their digital photographs? In *Proceedings of CHI '03*, pages 409–416, 2003.
16. B. Schneiderman, B.B. Bederson, and S.M. Drucker. Find that photo. *Communications of the ACM*, 49, 2006.
17. S. Scott, K. Grant, and R. Mandryk. System guidelines for co-located collaborative work on a tabletop display. In *Proceedings ECSCW 2003*, pages 159–178, 2003.
18. S.D. Scott, M.S.T. Carpendale, and K.M. Inkpen. Territoriality in collaborative tabletop workspaces. In *Proceedings of CSCW '04*, pages 294–303, 2004.
19. C. Shen, N.B. Lesh, F. Vernier, C. Forlines, and J. Frost. Sharing and building digital group histories. In *Proceedings CSCW '02*, pages 324–333, 2002.
20. L. von Ahn and L. Dabbish. Labeling images with a computer game. In *Proceedings of CHI '04*, 2004.
21. Andrew D. Wilson. Playanywhere: a compact interactive tabletop projection-vision system. In *Proceedings of the ACM UIST '05*, pages 83–92, 2005.