Balancing Awareness and Privacy in a Video Media Space Using Distortion Filtration

Michael Boyle
University of Calgary
2500 University Drive NW
Calgary AB T2N 1N4 CANADA
+1 403 220-3532
boylem@cpsc.ucalgary.ca

Saul Greenberg
University of Calgary
2500 University Drive NW
Calgary AB T2N 1N4 CANADA
+1 403 220-6087
saul@cpsc.ucalgary.ca

ABSTRACT
Collaboration among distributed workgroup members is hampered by the lack of good tools to support informal interactions. These tools either fail to provide teleawareness or enable smooth transitions into and out of informal interactions. Video media spaces—always-on video links—have been proposed as a solution to this problem. However, the "always-on" nature of video media spaces results in a conflict between the desire to provide awareness and the need to preserve privacy. The present study examines distortion filtration applied to always-on video as means of resolving this tension. Our discussions include the inter-related concepts of informal interactions, awareness, and privacy; and the treatment afforded by existing distributed collaboration support tools. We then outline the present study, where our goal is to understand the effect of distortion filtration on awareness and privacy.

Keywords
Tele-awareness, telepresence, privacy, informal interaction, video media spaces, distortion filtration, distributed collaboration.

1. INTRODUCTION
Collaboration may be loosely defined as people working together towards common goals, sharing ideas, and exchanging information. Humans have a flexible protocol for managing collaborative interactions that facilitates smooth and efficient collaboration, that is, collaboration in which content is communicated freely, correctly, and with a minimum of hassle. This protocol relies on a combination of subtle and explicit signals for expression including words, vocal intonations, gaze, posture, and gestures. Because much of this protocol was developed between people who are physically proximate to one another, it is no surprise that groups generally collaborate best when their members are co-located.

Instead of enjoying unfettered face-to-face contact—arguably the most natural way for humans to interact—distributed workgroups must rely on telecommunication media and computer-supported collaboration tools to work together. Cognitive psychologists suggest that for such a tool to effectively support collaborative activities, it must enable people to utilize the same social protocol for governing interaction as that which is used for face-to-face interactions [7]. While a plethora of tools for distributed collaboration has been developed, very few are capable of supporting people's natural social protocol.

In this paper, we will consider the design of a video-based media space that supports the way people naturally balance privacy and awareness. We begin in Section 2 by describing what we mean by informal interaction, and will include a summary of how technology tries to support this. In Section 3, we discuss how awareness of others is fundamental to supporting informal interaction, followed by how media spaces have been designed to support the two. Privacy concerns are presented in Section 4. Sections 5 and 6 describe our current approach. We introduce filtration as a means for balancing awareness and privacy, we discuss lessons learned by other researchers as well as ourselves, and we detail the study of the effects of filtration on awareness that we are currently running.

2. INFORMAL INTERACTIONS
2.1 Nature and value
At the heart of any collaborative process lie informal interactions. These are brief and spontaneous interpersonal exchanges where the participants, location, timing, and even agenda are not planned in advance [9]. Transitions into and out of informal interactions are made implicitly, and are thus the process is graceful, lightweight, and recoverable. Similarly, changes within informal interactions may also be made implicitly, making informal interactions smooth and adaptive. Lastly, in physically co-located situations, we find that informal interactions capitalize on a wide array of real-time communication channels (e.g., visual, aural, sensory), and thus are rich, interactive, and adaptive.

The unplanned nature of informal interactions is important. Because they are lightweight enough to allow participants to capitalize on serendipity and spontaneity, informal interactions occur very frequently. Because participants do not invest much energy in forethought and preplanning, they are free to adapt the nature of their interactions to best fit their immediate needs. Indeed, it has been claimed that informal interactions would not occur if they had to be planned [9]. In practice, the flexibility of informal interactions means they may be easily abandoned or postponed, that they can quickly shift between social and work contexts and between private and public matters, and that episodes of these interactions can effortlessly gain or lose participants. The length can be a few seconds but can easily stretch into hours. This level of freedom and degree of control over the particulars of interaction is rarely found in pre-planned interactions, which typically do not deviate much from their prescribed formats and agendas and, as a result, are often seen as slow and incomplete.

Although informal in tone, interactions of this type are of great importance to collaboration. Much valuable work-oriented information is exchanged by way of it. Furthermore, much of this

information is not exchanged by other means [9]. Thus many experts in groupware argue that a distributed collaboration environment's failure to support informal interactions will be to the detriment of its users.

2.2 Example systems supporting informal interactions

Communication tools designed to support informal interactions include examples such as the telephone, desktop videoconferencing (DVC), and experimental systems such as CRUISER [4,11]. These tools all follow a connect-interact-disconnect engagement model. This model works best for planned interactions and is advantageous because it reserves the highest quality of service for times of actual interaction—when it is needed most. However, this model isolates distributed group members when not directly engaged in interaction, and makes it harder for them to recognize opportunities for informal interaction.

Placing a call is heavyweight and requires explicit thought and action. Telephones mitigate this problem through ubiquity and fast connect times. DVC systems, where connections typically take 8 to 15 seconds to establish, cannot claim such an advantage. Worse, the incremental benefit to the outcome of collaborative tasks attained by adding video to an audio-only conference has been shown to be small [12,13,14]. This is not to say that video is without merit: there is strong evidence that adding video increases users' satisfaction with the collaborative process and the communication media.

The CRUISER system offered random connections to support serendipitous informal interactions (like bumping into a colleague in on one's way to the printer), but in practice was found to end up connecting strangers who had no interested talking with one another. Although serendipitous informal interactions are still possible in systems that use the telephone-model for engagement (e.g., when someone other than the intended person answers the call, or is present when the call takes place), the process is complicated by the fact that callers cannot tell if it is a good time to call unless they place the call and explicitly ask. Aside from being explicit, systems that use the telephone engagement model are intrusive. For example, in interviews conducted during observational studies of the CRUISER system, users reported that the video windows—which just popped up unexpectedly on their displays—were at times quite disruptive, much like the loud ring of an unexpected telephone call.

2.3 Summary

The key weakness found in all of the collaboration tools mentioned here is an inability to provide the teleawareness and sense of telepresence that are needed for informal interaction. They fail to provide enough informal awareness to infer presence and availability, or they do so in an awkward way. Next, we address teleawareness and the role it plays in supporting informal interactions.

3. AWARENESS FOR INFORMAL INTERACTIONS

3.1 Nature and value

Awareness is an unobtrusive background understanding of our shared work environment, the people in it, and their activities [7]. People incessantly accrue awareness information from subtle and implicit visual and auditory cues sensed and tabulated at the periphery of our consciousness. Rarely does one actively seek to acquire awareness.

Informal awareness is a sense of the presence, availability, and activities of others. It is particularly easy to acquire this kind of awareness when workgroup members are physically proximal to one another [7,9].

Conversational awareness is an understanding of the state of one's conversation. It is picked up from visual cues, such as eye contact, facial expressions, gestures and posture, and from auditory cues such as intonation and the use of particular words [7]. Conversational awareness is tightly coupled with informal interactions, and demands use of very rich and high quality channels such as those provided by face-to-face interactions.

Both informal and conversational awareness are critical to establishing and regulating informal interactions and facilitates transitions into, out of, and within informal interactions in many key ways. Informal awareness helps people stay on top of who is available or interruptible. Indeed, before interaction is engaged, this awareness helps participants decide many of the particulars of interaction. Knowing who is present or available helps one decide whom to contact, and this, in turn has implications for the choice of medium, locale, and timing. Conversational awareness is the bridge that lets people act upon informal awareness, turning opportunity into conversation.

Conversational awareness makes interactions smoother by helping to regulate the flow of interaction. It helps speakers and listeners exchange answers to mechanical and affective questions, so as to make incremental refinements to interaction [7]. By constantly incorporating small changes derived from conversational awareness cues, changes within the conversation—for example, speaker/listener role reversals, and shifts in topic—can be made in a less abrupt and less explicit fashion.

Awareness is not a one-sided affair: all participants maintain some awareness of all of these aspects. This mutual understanding facilitates co-operative decision-making and allows participants in informal interactions to quickly decide such things as: the entry or departure of a participants; changes of locale; and, whether to postpone further discussion and reschedule it for a better time.

3.2 Example systems supporting informal awareness

Teleawareness provided by existing support systems vary widely in fidelity. At the coarse end of the spectrum are instant messengers (IM) such as ICQ [8] and MSN Messenger [10]. These systems couple synchronous chat and lightweight asynchronous messaging with on-line presence information. One's on-line state is shown using an iconic representation, with states such as available, busy, away, and off-line. In practical experiences, even this crude sense of telepresence seems to be better than none, but there remains a need to further examine patterns of instant messenger use before firm conclusions may be drawn.

Portholes [3] and similar systems lie in the middle of the teleawareness information spectrum. These provide periodic video snapshots from a federation of remote sites. This low-bandwidth approach to providing long-term disengaged

teleawareness generated positive feedback from users in field studies. It enriched the sense of community both remotely across sites, and locally within each site.

At the high end of this spectrum are video media spaces: always-on video links, typically among offices or common spaces [1,5,13,14,15]. The fact that the audio and video channels are always left on is critical to the potential success of video media spaces. It allows users to employ the same social protocol governing co-located interactions to managing interactions in a distributed setting. Furthermore, it enables participants to maintain mutual awareness of each other during engagement transitions, and so these transitions are made smooth, graceful, and recoverable. Experiences with video media spaces [1,5,13,14] show that by supporting such lightweight and natural transitions into and out of engagement, they extend the effective interaction space.

3.3 Summary

Just as there are many sources of teleawareness, so too are there substantially different strategies for providing teleawareness. These differ in fidelity from instant messengers to video media spaces. It remains to be seen if any of these tools actually provide the telepresence needed, but these approaches all appear promising in the few systems implemented. However, in video-mediated teleawareness systems such as Portholes and video media spaces, there exist obvious concerns regarding privacy [1,5,13]. Next, we address the issue of privacy and the treatment it has received in existing teleawareness and informal interaction support tools.

4. PRIVACY

4.1 Nature and value

Privacy is generally taken to be the ability to manage what information about oneself is made public, to whom it is made available, and for what purpose it is used [7]. The ability to preserve one's privacy is often taken together with the right to ensure that any information about oneself is accurate and used fairly.

Solitude is a related concept often mistaken for privacy. Solitude is the ability to manage interruptions and distractions [7]. Since "interruptions" and "distractions" are felt to be negative concepts, it may be better to think of solitude as the ability to manage the "if, when, why, how and with whom" of engaging in interaction.

The preservation of privacy in co-located settings hinges on reciprocity in our collaborative environment [7]. Reciprocity is a simple rule regarding interaction: in order to see, one must be able to be seen, or in order to hear, one must be able to be heard. Reciprocity helps mitigate privacy concerns by discouraging eavesdropping or spying, by making it difficult to get away with such behaviors, and by letting people know that they are being watched so they can adjust their behaviors accordingly.

Awareness in co-located settings is reciprocal, and mutual awareness helps us preserve each other's solitude. When participants are aware of the availability of one another, they are better able to make the decision whether to engage in interaction. It should be clear that a distributed collaboration support tool that forces a user to turn it off in order to preserve any measure of privacy or solitude (like video media spaces or telephones) is poorly designed. Turning the system off defeats the system's purpose because it no longer allows any awareness to flow through.

4.2 Revisiting systems by privacy

Although widely known to be insecure, telephones are generally considered a comfortable tool for exchanging sensitive information because of its predominately dyadic character. This seems to defy the fact that it is known to be insecure, and may be the result of acclamation over long-term use. Furthermore, the telephone is also quite inadequate at preserving one's solitude, because it fails to provide availability awareness: the only way to know for sure if someone is available for interaction by telephone is to call him and interrupt him, and therefore invade his solitude. Screening calls is not entirely satisfactory because it cuts off all means to quickly pass urgent messages that supersede the need for solitude, and confuses the "absent" and "unavailable" states. As it was based on the telephone model for engaging in interaction, the CRUISER system suffered from similar problems related to intrusiveness. Surprisingly, most users were comfortable with the idea of strangers looking into their offices, and this is likely because the system enforced reciprocity.

Most instant messengers require username/password authentication, and use explicit (and typically reciprocal) authorization schemes for determining who will see one's on-line status and send one messages. Although IMs normally appear as unobtrusive icons, "nuisance" messages sent using them could be very obtrusive. As these kinds of frequent interruptions are possible in other communication media (e.g., telephone, face-to-face contact) it could be argued that such situations are best resolved by natural social protocol, such as by sending a curt "go away" message, or by making oneself more inaccessible.

While Portholes-like systems don't threaten solitude, they—along with video media spaces—have significant privacy issues. Privacy of outgoing content is a bigger concern than solitude violations [4] but most existing systems have only crude privacy support. There were facilities to edit the local snapshots taken by the Portholes system, but the interface implementation was hard to use and users easily confused the absence of fresh snapshots for "technical difficulties." Some media spaces support disconnected states, but, as when screening telephone calls or closing one's office door, such behaviors confuse unavailability with absence. Other media space designs ignore the privacy issue altogether, but observations of actual use show that even among intimate collaborators privacy-awareness tensions exist [1].

4.3 Summary

There has been, historically, poor support for preserving privacy and protecting solitude in distributed collaboration support tools. Privacy is itself a nebulous concept, highly dependent on context. Consequently, most designers have taken a lazy approach to the subject. While many raise it as a concern, few address it. This is a serious omission, for we expect that privacy over the Internet will be a key issue for the 21st century. Ideally, one would like to capitalize on a video media space's ability to provide awareness and lightweight transitions into and out of interaction, but balance these aspects against a real need to provide control over privacy and solitude.

5. DISTORTION FILTRATION

5.1 Nature and value

Distortion filtration is proposed as a possible means for striking a balance between the desire to provide teleawareness and the need to preserve privacy. It is the process of algorithmically manipulating the contents of the audio and video streams to selectively obscure various levels of detail in the sound or picture. The goal is to hide elements of the picture or sound that might be deemed sensitive —details which one wishes to keep private— while still giving remote participants the ability to extract just enough awareness and presence information to make confident decisions regarding one's availability and interruptability.

Many filtration effects are possible: pixelization; smoothing; convolutions (e.g., Sobel, Laplace operations); overlays; and, translatory distortions (e.g., wave, ripple, fish-eye) are but a handful. Several examples are illustrated in Figure 1. These can be also applied across a continuum of extents (i.e., levels), as illustrated in Figure 2. Furthermore, the filter may be coupled with an excitory "lens" which determines the level to filter at a given point in a given image in the sequence. Figure 3 gives two example excitory lenses. The activity-selective lens applies the filter if a pixel in the current frame differs substantially from the corresponding pixel in a reference frame—typically a snapshot of the scene background, taken when no one is present. The second example, a radial lens, applies the filter in ever increasing amounts outward from the center of the image.

Figure 1. Example distortion filters: (clockwise from top-left) undistorted, pixelize, wave, and smoothing.

5.2 Existing systems

The visual channel, as suggested earlier, is an important part of interactions, but as demonstrated by the success of the telephone, it is by no means a requirement. Hudson and Smith [12] developed an audio media space (common audio channel shared among many participants) that does not disturb or annoy participants with unwanted audio fragments. Their report describes a system that provided both a background ambient audio signal, representing the overall audio activity level at a number of distant sites, as well as facilities to "tune in" to a specific speaker.

Figure 2. Pixelize filter at four different levels.

Figure 3. Pixelize filter used with two different excitory lenses: (left to right) activity selective, and radial.

Zhao and Stasko [15] applied a number of distortion filters to video clips, and then tested volunteers for accuracy in recognizing actors and activities. Filtration was performed at a single level. The filters examined were: pixelization; Sobel; a "shadow" filter (a pixelize filter coupled to a lens similar to the activity-selective lens discussed earlier, only that the filtration is applied to the reference frame); and a "live shadow" filter (an activity-selective lens). Twenty volunteers, mostly unfamiliar with the actors in the scenes shown, were used in tests. After viewing one of 21 video clips in one two sizes, a volunteer was asked specific questions about the location viewed, the actors in it, and their activities. The questions asked could be deemed "loaded", and the only metric recorded was the number of correct responses. Informal interviews were used to elicit qualitative information about usage and users' reactions, and the feedback given was incorporated into the design of an improved filter, the results of which have not yet been reported.

There are real-time performance issues when adding distortion filtration to live video links. A typical DVC application must

capture, compress, transmit, receive, decompress and render video frames many times a second. Common desktop hardware/software/network configurations already have a difficult time keeping up with these core tasks, even at low frame rates, small frame sizes, and using highly lossy compression techniques. There is not much time left for sophisticated or computationally intensive filtering algorithms without the benefit of specialized hardware. Consequently, filtering algorithms toyed with thus far have been designed without much regard to human physiology and perception. Moreover, even if one is lucky enough to get an algorithm to work within real-time constraints, there does not yet exist a systematic way of testing a filter for its effects on awareness and its effectiveness in preserving privacy.

5.3 Summary

The idea of using distortion filtration in a video media space seems, at least superficially, like a good approach to balancing the tension that exists between maintaining awareness and preserving privacy. However, numerous questions remain unresolved by the work already done in this area. What is the impact of a filter on one's ability to extract awareness cues from a video scene? What are the characteristics of an effective privacy-preserving distortion filter? To what range of extents should a particular filter be applied? How should the extent of filtration across its range of useful values be controlled?

6. PRESENT STUDY

We are in the process of constructing always-on video links for awareness and interaction that reflect the lessons learned from previous experience. To this end, the present study will examine the impact the pixelizing and smoothing filters have at each of ten different extents on an observer's ability to extract key awareness information from a filtered video sequences These filters were chosen because the can be performed within real-time constraints.

Observers will view two video scenes; each scene will be filtered using either the pixelizing or the smoothing filter. Scenes will use zero or more actors. Actors will be engaged in activities such as: eating; talking on the telephone; working alone; or, engaged in conversation. Typical work locales will be featured, such as private offices, coffee rooms, corridors, laboratories, conference rooms, and home offices.

The testing protocol is as follows. The observer views the video sequence with the filter applied across its entire range of extents, starting at maximum distortion. He answers a set of open and closed questions about the location, the actors in it (if any), their activities, and their state (e.g., disposition). For each question, the observer is asked to mention how it is he came up with his answer (i.e., from which features in the image sequence does he infer his answer), and how confident he is that his answer is correct. When the questions for one filtration level are complete, the level of filtration is reduced and questioning repeats. The process continues until the video becomes undistorted.

To elicit qualitative information about the observed filter's capacity to preserve one's privacy, at the conclusion of the tests for one of the filters, we ask the volunteer to imagine himself as the actor in the scene viewed, performing the same activities. When then ask him to choose the level of filtration at which he feels comfortable letting certain classes of people (e.g., strangers, close friends, superiors) view. The conclusion of an entire test session will be marked by an informal interview with open-ended questions about filtration and privacy.

As the study is just underway, results are few. Results from pilot tests clearly show that a single extent of filtration is insufficient to cover all privacy-awareness concerns: filtration must adapt to the demands of circumstance. Ostensibly, this makes sense: in face-to-face contact, the resolution at which we see another person and the understanding of their availability we derive from looking at him improves as we near him.

Additionally, it has become clear in pilot testing that motion in video is a critical source of information about presence and activity. Observers have great capacity to infer availability from subjective assessment of the motion they see. Effective filters should not destroy the continuity of motion throughout a video sequence, particularly if the underlying video is at a low frame rate, where changes resulting from rapid motion can be dramatic.

Pilot-test volunteers responded that they felt that the filters tested would be effective at preserving privacy, but many still wonder why one wouldn't turn the system off entirely in such sensitive situations.

7. NEXT STEPS

Contingent upon the results of the present study, we hope to incorporate a privacy-preserving distortion filter into a teleawareness support tool. We next intend to examine lightweight techniques for controlling the level of filtration, such as reacting to information accrued from physical proximity detectors, motion recognition and other sources of teleawareness. Other issues to be examined include how to handle privatization/republications of sub-conversations in multiparty video media spaces, and how to smoothly integrate with other communication media and with workspace tools.

At present, we have made a simple networked, two-party video media space application that incorporates the pixelize filter at three different qualities of service: high frame rate, no distortion; high frame right, moderate distortion; low frame, high distortion. Figure 4 shows the highest and lowest qualities of service offered.

Figure 4. Highest and lowest qualities of service offered in our video media space app. Utilizes the pixelize filter.

This program demonstrates many important principles learned from previous research. First, it is reciprocal: the quality of service viewed by one party is the same as that viewed by the other. Second, quality of service is implicitly controlled. We suppose that placing the mouse pointer inside the video window is an implicit cue that one is engaged in interactivity with the remote party, and so higher qualities of service should be used. We do

not know if this is an appropriate technique: although possibly interesting, it is largely intended as a placeholder for some other implicit control mechanism.

Third, changes in quality of service are mutually decided. One party may raise or lower the quality of service one level, but both parties must cooperate to take it to either extreme. Table 1 describes the behaviour of the quality of service control featured.

Pointer positions	Frame rate	Filter level
Both inside	High	No distortion
Inside and outside	High	Moderate distortion
Both outside	Low	High distortion

Table 1. Mapping of pointer positions (relative to each party's video window) to quality of service factors.

To discourage turning the system off to assure total privacy, the system allows one to "block" his video. While one party blocks his video, he sees the image of the back of a hand superimposed over the remote party's video image; the remote party instead sees the image of the palm of a hand. Both parties may block at any time—indeed even at the same time—but each is responsible for removing the block that he raises. We have attempted to give blocking a tangible aspect: it is toggled on or off by using one's hand to cover the camera so that one's image goes black for a few frames. We have chosen this approach to closely resemble one's tendency to reach for the camera in "dire" circumstances. Figure 5 shows the blocked video, as seen by the blocking and blocked parties.

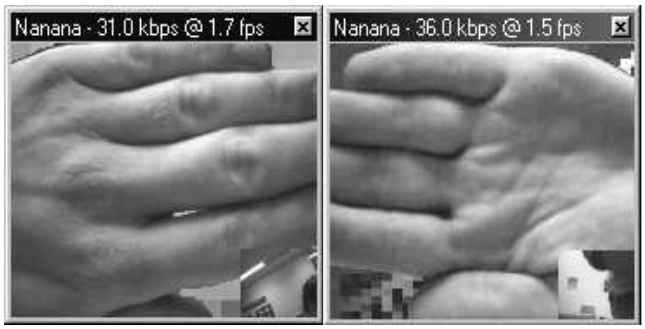

Figure 5. Blocked video as seen by blocking (left) and blocked (right) parties. Note video in corners can still pass through.

This video media space application is still quite experimental. The decision to use the pixelize filter was made arbitrarily, and will be revisited when the present study completes. The implicit and explicit interaction techniques have not undergone user testing. Many important technical problems relating to the capture, compression, and transmission of live video have been resolved and we are presently able to rapidly develop video-based applications using an iterative approach.

Lastly, more work is also needed in the field of video-mediated telepresence and teleawareness. It is not yet known conclusively if video provides any useful teleawareness or telepresence cues over and above other sources (e.g., on-line presence). Moreover, the minimum quality of service (e.g., frame size, rate, latency) needed to provide telepresence, and the maximum useful quality of service are still not known. In our present course of research, we hope to turn our attention to some of these issues.

8. ACKNOWLEDGEMENTS

Funding for this research provided by the National Science and Engineering Research Council, and Microsoft Research.

9. REFERENCES

[1] W. Buxton. Living in augmented reality: Ubiquitous media and reactive environments. In *Video Mediated Communication*, Hillsdale, NJ, 1997.

[2] A. Cockburn and S. Greenberg. Making contact: Getting the group communicating with groupware. In *Proceedings of the ACM/SIGOIS Conference on Organization Computing Systems (COOCS'93),* pages 31–40, California, 1993.

[3] P. Dourish and S. Bly. Portholes: Supporting awareness in a distributed work group. In *Proceedings of the ACM/SIGCHI Conference on Human Factors in Computing Systems (CHI'92),* pages 541–547, Monteray, CA, May 1992.

[4] R. S. Fish, R. E. Kraut, R. E. Rice, and R. W. Root. Evaluating video as a technology for informal communication. In *Proceedings of the ACM/SIGCHI Conference on Human Factors in Computing Systems (CHI'92),* pages 27–34, Monteray, CA, May 1992.

[5] W. Gaver, T. Moran, A. MacLean, L. Lövstrand, P. Dourish, K. Carter, and W. Buxton. Realizing a video environment: EuroPARC's RAVE system. In *Proceedings of the ACM/SIGCHI Conference on Human Factors in Computing Systems (CHI'92),* pages 27-34, Monteray, CA, May 1992.

[6] S. Greenberg. Peepholes: Low cost awareness of one's community. In *Proceedings of the ACM/SIGCHI Conference on Human Factors in Computing Systems (CHI'96),* pages 206–207, Vancouver, BC, April 1996.

[7] C. Gutwin. Workspace awareness in real-time distributed groupware. PhD thesis, Department of Computer Science, University of Calgary, pages 11–31, 1998.

[8] ICQ Inc., *ICQ Version 99b Version b.3.19,* ICQ Inc., 1999.

[9] R. Kraut C. Edigo, and J. Galegher. Patterns of contact and communication in scientific research collaboration. In *Proceedings of the Conference on Computer Supported Cooperative Work (CSCW'88),* pages 1–12, New York, NY, 1998.

[10] Microsoft Corp. *MSN Messenger,* Microsoft Corp., 1999.

[11] R. W. Root. Design of a multimedia vehicle for social browsing. In *Proceedings of the Conference on*

Computer Supported Cooperative Work (CSCW'88), pages 25–38, New York, NY, 1988.

[12] I. Smith and S. E. Hudson. Low disturbance audio for awareness and privacy in media space applications. Technical report, Graphics, Visualization, and Usability Center, Georgia Institute of Technology, Atlanta, GA, 1995.

[13] J. C. Tang, E. A. Isaacs, and M. Rua. Supporting distributed groups with a Montage of lightweight interactions. In *Proceedings of the Conference on Computer Supported Cooperative Work (CSCW'94),* pages 23–34, Chapel Hill, NC, October 1994.

[14] J. C. Tang and M. Rua. Montage: Providing teleproximity for distributed groups. In *Proceedings of the ACM/SIGCHI Conference on Human Factors in Computing Systems (CHI'94),* pages 37–43, Boston, MA, April 1994.

[15] Q. A. Zhao and J. T. Stasko. Evaluating Image Filtering Based Techniques in Media Space Applications. *In Proceedings of the Conference on Computer Supported Cooperative Work (CSCW'98),* pages 11–18, Seattle, WA, 1998.

Designing an Integrated Bookmark / History System for Web Browsing

Shaun Kaasten
Department of Computer Science
University of Calgary
Calgary, Alberta, CANADA T2N 1N4
Tel: +1-403-220-7686
kaasten@cpsc.ucalgary.ca

Saul Greenberg
Department of Computer Science
University of Calgary
Calgary, Alberta, CANADA T2N 1N4
Tel: +1-403-220-6087
saul@cpsc.ucalgary.ca

ABSTRACT

Current commercial web browsers such as Netscape Navigator and Microsoft Internet Explorer attempt to make it easier for users to return to previously visited web pages. They offer three separate but important facilities: the back button, a bookmark system, and a history list. However, research indicates that users are not utilizing all of these systems effectively. In this paper, we present a single integrated history that unifies functionality similar to the back button, bookmarks and history lists. We also show how these mechanisms can be improved upon. We pay special attention to the ways pages are presented within a display, and on providing lightweight means for marking, recalling, and revisiting pages. The prototype we are developing works within Microsoft Internet Explorer. It presents a history list that represents pages by visual thumbnails as well as titles and URLs. It incorporates bookmark functions through a dog-ear metaphor, and it contains methods to filter a history list. Finally, it is tightly linked with the back/forward buttons.

Keywords

History, page revisitation, reuse, navigation, browser design, world wide web, hypertext.

1. INTRODUCTION

As the Internet continues to grow, the World Wide Web is becoming a mainstream information medium. People are relying on it more than ever for education, business and entertainment purposes. However, the World Wide Web is both enormous and largely unorganized, making it difficult to navigate effectively.

We all know how difficult it can be to find specific information on the web. One solution is better Internet Search Engines. This has become so important that the major players on the Web are in a race to provide effective search engines that can find what people ask of them, where the winner will get the most people and the most advertisers.

From the user's perspective, there is another problem that is getting less attention: how to return to information they have previously found. While a search engine might find the desired needle in the web's haystack, finding the same needle in the haystack later often requires them to search or navigate again, often with laborious efforts. Relocating a web page viewed days, weeks or even hours ago can be a frustrating and fruitless experience. Of course, current commercial web browsers such as Netscape Navigator and Microsoft Internet Explorer are not blind to people's need to return to previously visited pages. Each provides almost identical functions for returning to these pages, with the major ones being `Back`, the history list, and bookmarks.

In theory, these mechanisms should be heavily used, for almost 60% of all pages a person visits are to ones that they had seen previously [9]. Yet research indicates several problems with these mechanisms. While `Back` is heavily used, people have an incorrect model of how it works, which leads to surprises when just-visited pages are no longer reachable through it [6, 7]. Bookmark and history systems are used infrequently in practice [1, 9]. We believe that one of the reasons for these problems is that browsers provide revisitation systems in a fragmented, un-integrated, and heavyweight manner. `Back`, history and bookmarks all use dissimilar underlying models, different interfaces, and various ways of sorting and presenting groups of candidate pages. Recognizing pages in history lists can be problematic. Remembering to bookmark a page and managing the clutter of pages in a bookmark list can be onerous.

In this research, our goal is to integrate the idea of `Back`, history and bookmarks into a single integrated revisitation system that captures the best features of existing systems while remedying their known deficiencies.

2. CURRENT SYSTEMS

2.1 Bookmarks

Bookmarks have been an important feature, included in even very early web browsers. The concept is simple – when a user finds a page worth returning to, he or she asks the web browser to create a bookmark for it. Some time later, the user selects the bookmark to get back to the particular page. Unfortunately, the system fails in practice. While users are consistently return to previously visited pages, they rarely do this through bookmarks. The Tauscher and Greenberg study found that bookmark navigations only made up 2% of all navigation [9]. This can be traced to some major shortcomings of bookmark systems.

First, bookmarks are inefficient. A study on bookmark usage found that users continually accumulate bookmarks, which remain long after they are actually needed [1]. Thus, the bookmark collection becomes cluttered with unimportant items that inhibit the user from seeing the newer, now important bookmarks. In an attempt to reduce this problem, bookmark systems allow the creation of folders for organization. However, the study found that most users put off organization until it becomes absolutely

necessary. Quite Simply, people are unwilling to devote the time needed to keep bookmarks organized.

Second, bookmarks are based on the notion that people know which pages they will need to return to. This requires a person to immediately determine the value of a page. However, circumstances can change, and a page that one decided to skip through last week can suddenly be valuable today. Users are poor predictors of what will be important at a later time.

2.2 History Lists

Web browser designers have recognized that people need a way to return to previously visited sites that have not been explicitly bookmarked. All contemporary browsers offer a fairly similar solution – the history list. A history list presents the user with a sorted list of every web page visited. The list can be sorted by the page title, URL address, the date of last visit, and even by number of times visited.

Like bookmarks however, history lists break down in real-world use. The Tauscher and Greenberg study found that users do not take advantage of the history system, as it accounted for only 1% of web page navigation [9]. Like bookmarks, this finding is related to some shortcomings of current history systems.

Figure 1 – Our prototype history revisitation system

The history list in early browsers had a great weakness- it was largely hidden from the user. Raising the history list required the user to select an option from a rarely used menu. While some users were not even aware of the option's existence, others found the process of uncovering it from the menu very inconvenient. As well, the history list was in a separate window that incurred the added work of window management. Current browsers have attempted to bring the history list closer to the user. Microsoft's Internet Explorer 5, for example, provides a large history button along side the more familiar controls, encouraging the curious user to try it out (see top-right of Figure 1). It also includes it as a tiled window, called an 'explorer bar' (similar to the left panel in Figure 1), which lives within the main browsing window.

When a user does uncover the system, he or she must then try to find the desired page within the list. This could be problematic.

To begin, web pages are often represented by title. Studies have pointed out how web page titles are frequently meaningless to the user [4]. First, it is the web page author, not the user who creates it. The author-supplied title may not match how a user thinks of that page. Second the web page author often gets it wrong: they may forget to include a title, or they may give a page the same title they had given to other pages within a site. This could be by accident (e.g., as when a titled page template is used over and over again and the author forgets to change it), or by intent (e.g., a title like 'FAQ' may identify a site full of 'FAQ's). When the user is confronted with a list of such unreliable titles, it can be virtually impossible to recognize the needed page.

Web pages can also be identified by their URL address. This address allows an individual web page to be instantly accessed from anywhere in the world. While these addresses are globally unique, they are often long and obscure. Users cannot be expected to memorize most URLs, and even recognizing a page from its URL can be difficult. In the best cases, the URL offers clues about a web page, especially when it contains a recognizable name such as Microsoft. Even here, the small display space given to the history list may mean that the entire URL is not visible to the user. In the worst case, the URL is meaningless (such as those produced automatically by databases). Thus, a history list of URLs is, at best, a hit and miss proposition for most users [4].

3. PROTOTYPE SYSTEM

We are currently developing a prototype system that works within Microsoft Internet Explorer. Some of our ideas were first explored in earlier versions [4, 5]. Figure 1 illustrates one of our prototypes. The unifying principle of this prototype is that all visited pages are organized as a recency-ordered history list, with duplicate pages shown only in their latest position.

3.1 Back Button Integration

Our first challenge was to integrate the back and forward buttons with the history list. All commercial browsers implement a stack

rather than a recency ordered list. We altered the back and forward buttons so that they work on a recency-based model instead of a stack: the buttons are thus just shortcuts for moving up and down the history list [4]. By bolding the currently visited item on the list, users see immediate navigational feedback on their `Back` and `Forward` actions. Aside from the fact that `Back` and `Forward` are now integrated with history, there are several advantages [7]. First, a person can always return to a page with a recency-ordered back button because, unlike stack, pages are not pruned from the list. Second, people who were asked how the conventional stack-based back button works all had the incorrect model that it was based on a recency ordered list [6]. Thus our new back button reflects people's model of how they *think* `Back` should work.

3.2 Implicit / Explicit Bookmarks

Our next challenge was to integrate bookmarks into the history list. We have transformed bookmarks into two different features.

First, *implicit bookmarks* are pages that are visited frequently: these pages are automatically visually distinguished from pages visited rarely. Both Figure 1 and the small inset here shows how we mark these pages with a vertical red band, where the height of the band indicates the degree of the page's visit frequency. If there is no band, the page has been visited only once or twice. Modest sized bands often indicate a 'hub' page i.e., a page that contains many children pages that have been visited. Since a user typically returns to the hub page to chase the next link [9], these typically show a higher number of hits. Large bands often indicate important pages, such as home pages for frequently visited sites.

Second, *explicit bookmarks* are just pages on the history list that are tagged as special: we call these *dog-ears*. A user can quickly set a dog-ear by right-clicking over a page on the history list and selecting 'dog-ear' from a popup menu; users then have the opportunity to change the page's title if they wish. Dog-eared pages look different from non-dog-eared pages. For example, in Figure 1 (as well as this inset) we see that the 'Grouplab' and the second 'University of Calgary' page are dog-ears, as illustrated by the folded down corners in the upper right corners of the images.

3.3 Search Filters

To make this all work smoothly, users can perform dynamic queries [8] on the history list. By moving a slider (illustrated at the bottom of Figure 2), they rapidly and continuously filter the history list. As the slider is moved to the right, pages with low visit frequencies are immediately filtered from the list. When it is almost all to the right, only high frequency pages and dog-ears are visible. When positioned at the extreme right, only dog-eared pages are displayed: this transforms the general history list into a recency-ordered bookmark list.

We also have added a *search filter*, where the history list is filtered to display only those pages whose title contains the specified sub-string. As before, this is done through a dynamic query: as letters are typed, the list is immediately filtered to show only those matching pages.