Time Quilt: Scaling up Zoomable Photo Browsers for Large, Unstructured Photo Collections

David F. Huynh[†], Steven M. Drucker[‡], Patrick Baudisch[‡], Curtis Wong[‡]
[†]MIT Computer Science and Artificial Intelligence Laboratory,
The Stata Center, Building 32, 32 Vassar Street, Cambridge, MA 02139, USA
[‡]Microsoft Research, One Microsoft Way, Redmond, WA 98052, USA
dfhuynh@csail.mit.edu, {sdrucker, baudisch, wong}@microsoft.com

ABSTRACT

In the absence of manual organization of large digital photo collections, the photos' visual content and creation dates can help support time-based visual search tasks. Current zoomable photo browsers are designed to support visual searches by maximizing screenspace usage. However, their space-filling layouts fail to convey temporal order effectively. We propose a novel layout called time quilt that trades off screenspace usage for better presentation of temporal order. In an experimental comparison of space-filling, linear timeline, and time quilt layouts, participants carried out the task of finding photos in their personal photo collections averaging 4,000 items. They performed 45% faster on time quilt.

Furthermore, while current zoomable photo browsers are designed for visual searches, this support does not scale to thousands of photos: individual thumbnails become less informative as they grow smaller. We found a subjective preference for the use of representative photos to provide an overview for visual searches in place of the diminishing thumbnails.

Categories & Subject Descriptors: H5.2 [Information interfaces and presentation]: User Interfaces. - Graphical user interfaces.

General Terms: Human Factors; Design.

Keywords: Digital photography; space filling; timeline; representative thumbnail; zoomable UI; semantic zooming.

INTRODUCTION

IDC's 2003 U.S. Consumer Digital Imaging Survey reported that 17% of its respondents took 50–100 photos per month, or equivalently, 3,000–6,000 images over a 5 year period. The sizes of these collections make it difficult to manually organize them to facilitate effective browsing and searching at a later time. Frohlich et. al. [6] reported that very few families participating in their user study systematically organized their digital photo collections.

Many users have already given up organization in other information domains such as email. However, in such textual domains as email, automatic indexing technologies can be called upon to support keyword searches effectively. For a visual domain such as digital photos, contemporary indexing technologies are still immature. Rodden and Wood [15] reported that the content-based image retrieval feature offered in their user study's test system was rarely used and its perceived utility was low.

In the absence of manual organization performed by the user, there is still the data collected by the digital camera, specifically, the visual content of the photo and the creation date. Not only is the creation date recorded automatically, it is also an essential factor by which people browse their photos [7]. With only creation date and visual content, people can still perform *time-based visual search* tasks such as:

- reminiscing over a past period of time, to answer such questions like, 'what pictures do we have from the last five years?'; and
- finding photos associated with some memorable events (e.g., 'that photo of you hanging from a cliff that I took on our trip to the Grand Canyon two years ago').

Zoomable photo browsers (e.g., PhotoMesa [2]) are designed for *visual searches* by maximizing screenspace usage, thus requiring minimal panning. However, by adopting a space-filling layout (e.g., quantum treemap [2]), they don't convey temporal order well: although they allow users to cluster photos by creation date, the clusters are not laid out in such a way that communicates their order in time.

Contributions

In this paper, we propose 2 modifications to the existing zoomable photo browsers that let them support *time-based* visual search tasks more effectively and over larger collections: a layout called *time quilt* (Figure 1) that makes a compromise between filling screen space and conveying temporal order; and the use of representative thumbnails to implement semantic zooming for better overview of larger collections.

RELATED WORK

Photo Browsing and Visualization – In 1999, Kuchinsky et. al. [9] introduced *FotoFile*, a consumer-oriented multimedia organization and retrieval system designed to support ex-

Copyright is held by the author/owner(s).
CHI 2005, April 2–7, 2004, Portland, Oregon, USA.
ACM 1-59593-002-7/05/0004.

ploratory browsing to address consumers' lack of economic incentives for annotating their photos to facilitate keyword-based searches. Also in 1999, Combs and Bederson [3] tested the first zoomable image browser on 225 images. Bederson followed up in 2001 with *PhotoMesa* [2], a second zoomable browser that used space-filling layouts. Kang and Shneiderman [8] introduced *PhotoFinder* in 2000, equipped with visual Booleans and dynamic query interfaces, designed primarily to address searching rather than browsing.

In 2002, Platt et. al. [13] performed one of the first formal user studies on *personal* photo collections (averaging 850 images). Their photo browser, PhotoTOC, used representative photos to show a table-of-content summary of photos clustered by date. From a six month-long study in 2003, Rodden and Wood [15] concluded that two of the most important features to support in photo browsers were (1) sorting by chronological order and (2) displaying a large number of thumbnails at once. Their participants most commonly wanted to browse their personal photos by event rather than by querying them based on more specific properties.

In 2004, Drucker et. al. [4] presented a careful selection of many previous research concepts integrated into a single browser, the *MediaBrowser*. By integrating temporal clustering with rapid selection, they were able to make it easier for users to annotate their collection. In an informal study, they found that loading *MediaBrowser* with more than 500 or 600 objects rendered individual thumbnails hard to distinguish by eye, thus identifying scalability as a topic for future work.

Time Visualization – There is much prior work on presenting temporal information. Ringel et. al. [14] investigated a timeline layout to present search results from personal information corpora. Mackinlay et. al. [10] developed the *Spiral Calendar* for rapid access to an individual's daily schedule by integrating details and context using a 3D spiral layout. *LifeLines* [12] is a tailorable environment for showing personal histories in multiple facets. *Lifestreams* [5] is a system for showing a user's personal file system in a timeline format. Limited support for time-based browsing of personal photo collections exist in commercial software.

Semantic Zooming – Semantic zooming was introduced by Perlin and Fox [11] in the *Pad* system in 1993 and later supported pervasively in the *Pad*++ system by Bederson and Hollan [1].

To address the problem of shrinking thumbnails, Suh et. al. [16] proposed a method of automatically cropping a photo to keep only its most salient region. This solution generates more recognizable thumbnails only to some relatively small size.

THE TIME QUILT LAYOUT

Space-filling layouts such as that used in PhotoMesa are designed to arrange several clusters of photos together such that:

- the white space between the photos is minimized; and
- the whole visualization exhibits a specified overall aspect ratio and the whitespace external to the visualization is minimized when zoomed out.

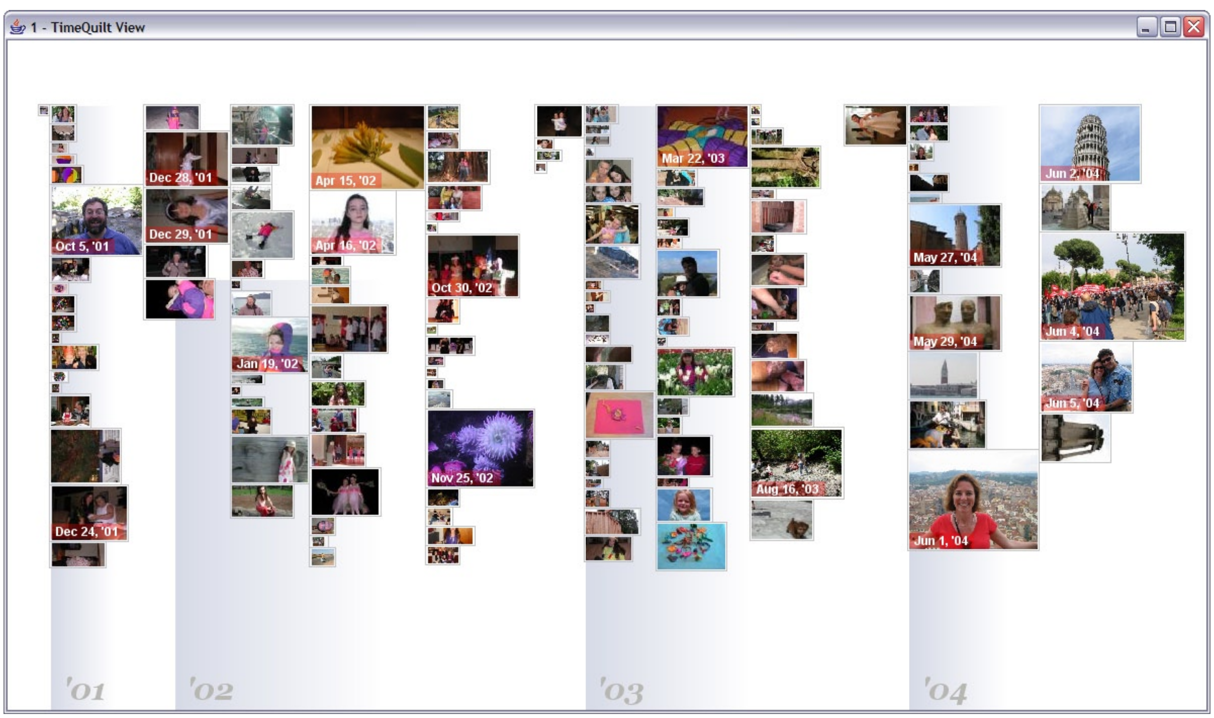
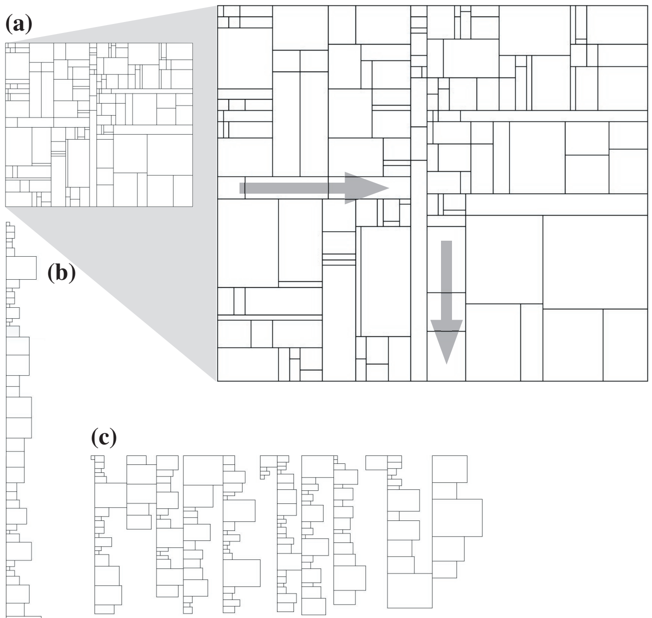


Figure 1. Time quilt – a layout designed to convey temporal order while making better use of screenspace than a timeline, showing approximately 5,500 photos with representative thumbnail overview



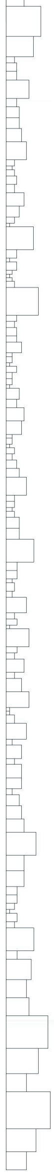


Figure 2. A space-filling layout (a) saves space but does not convey time; a timeline (b) wastes space and requires excessive panning due to severe aspect ratio; weaving the timeline along both dimensions (c) conveys time and achieves reasonable aspect ratio.

Such maximal use of screenspace supports visual searches effectively but it has 2 side-effects undesirable for conveying temporal order (Figure 2a):

- some clusters are elongated with severe aspect ratios that make it hard to fit them on the screen when they are zoomed in; and
- clusters consecutive in time are laid out sequentially sometimes in the vertical direction and sometimes in the horizontal direction, making it hard to find and pan through consecutive clusters (arrows in Figure 2a).

Furthermore, as the collection grows, the clusters are shuffled about in order to maintain the overall aspect ratio of the whole visualization. This display instability can affect the user's spatial memory of the visualization.

In a conventional timeline (Figure 2b), clusters can be shaped independently of one another, each exhibiting a reasonable aspect ratio. Clusters consecutive in time follow one another in a consistent direction. The overall visualization does not shuffle as the collection grows. However, a timeline exhibits a severe overall aspect ratio, thus making poor use of screenspace and requiring excessive panning when zoomed out.

We propose "weaving" the straight timeline to achieve better overall aspect ratio while retaining the timeline's ability to convey temporal order. Figure 1 shows one particular "weaving" layout called *time quilt*: clusters of photos are wrapped in vertical columns (similar to how text is wrapped in horizontal lines). The layout algorithm works as follows:

- All photos are divided by their creation dates into clusters using Platt's adaptive clustering algorithm [13].
- All photos are scaled down to thumbnails of the same size.
- The thumbnails of the photos of each cluster are laid out into a grid such that the overall aspect ratio of the grid matches that of the screen.
- The grids of thumbnails are ordered by time and then wrapped into vertical columns of some maximum height (e.g., 7,000 pixels). Column breaks are also inserted at year boundaries.

REPRESENTATIVE THUMBNAILS

Presenting photos as thumbnails works only for some number of images, beyond which the thumbnails become too small to carry information individually. We propose the use of representative photos to implement semantic zooming in zoomable photo browsers: when individual thumbnails become too small to be recognizable at a certain zoom level, a representative photo from each cluster is rendered in place of the diminishing thumbnails. The representative photo is scaled and cropped to best occupy the same space as the original cluster grid. While the use of representative thumbnails is independent of the layout, they are more useful when the aspect ratios of individual clusters are reasonable, as in either the timeline layout or the time quilt layout. In our prototype browser, we simply choose the middle photo of each cluster as its representative. Platt et. al. have suggested better ways for selecting representative photos [13].

USER STUDY

We conducted a user study to compare 3 layouts: a space-filling layout (quantum treemap [2]), the timeline layout, and the time quilt layout. We implemented 3 interfaces using the 3 layouts, as well as a fourth interface using the space-filling layout but without representative thumbnails. All interfaces supported zooming through the mouse-wheel and panning through left-button dragging. Selection of a photo could be performed by left-clicking.

The test computer was connected to 2 monitors: the left one showed a target photo to be found while the right one showed one of the four interfaces. The participants' task was to browse the interface to locate the target photo and select it. Each participant was instructed to "go for speed" while maintaining reasonable accuracy. If a participant was unable to find a photo for several minutes, s/he was instructed to abandon the task.

10 people (2 females) participated in the study. They responded to a user study advertisement on the Photography

Enthusiast forum within our company's intranet. Before their user study sessions, the participants shared with us their photo collections ranging from 2,863 to 5,708 photos (median ~ mean = 3,994, stdev = 928). Each participant also selected 28 favorite photos from his/her collection, almost all from different events. We randomly divided these favorite photos into 4 groups to test the 4 interfaces, each group consisted of 2 training photos and 5 photos for the actual test.

We used a within-subject experimental design: each subject carried out the task 5 times (for 5 different target photos) on each of the 4 interfaces. In order to avoid sequence effects, the interface order was counterbalanced between subjects. Each participant received verbal instruction when the study session started and an introduction to each interface before the tasks on that interface were performed. After the introduction, s/he was encouraged to explore the interface by finding 2 training photos.

We had 2 hypotheses:

- Participants would complete the task faster in time quilt interface than in the other two.
- Participants would complete the task in the space-filling interface faster with rather than without representative photos.

The participants achieved better task completion time on time quilt than on timeline and space-filling (44.2% and 45.1% faster respectively, one-way ANOVA test $p = 0.002$, $F(3,173) = 5.314$). Note that we did not count 14 trials (out of 200) in which the task completion time exceeding 3 minutes. In such trials, participants did not recall enough information to locate the target photos and simply resorted to panning over the entire visualizations.

The participants also specified that they "preferred the use of representative photos" (Likert scale of 7: M = 5.70, SD = 1.636, t-test = 0.009) even though they did *not* find that "the representative photos were accurate representation of the clusters" (Likert: M = 4.30, SD = 1.418, t-test = 0.520).

DISCUSSION

Although every digital photo always carries a creation date, the dates of the images shared by the participants turned out not entirely reliable. Some participants included scanned images in their photo collections and in the 28 favorites that they selected; the dates on these images were not the dates of the corresponding events. Also, some participants' cameras had faulty dates: some pictures taken in 2004 were dated 2003. In addition, participants included photos taken by someone else. As a result, they found it hard to remember approximately when those photos were taken.

As future work, there is much to be done. The issues to tackle include the followings:

- Better selection of representative thumbnails should be investigated as well as their use in the visualization. For example, many representative thumbnails can be shown instead of just one per cluster. Several levels of semantic zooming can also be explored.
- Other ways of weaving the time line should be tried. As well, better zooming and panning techniques should be incorporated in the comparisons.
- Strategies for dealing with inaccurate or missing dates and for taking advantage of any user provided metadata should be explored.

REFERENCES

[1] Bederson, B.B. and J.D. Hollan. Pad++: A Zooming Graphical Interface for Exploring Alternate Interface Physics. *Proc. UIST 1994*, 17–26.

[2] Bederson, B. PhotoMesa: a zoomable image browser using quantum treemaps and bubblemaps. *Proc. UIST 2001*, 71–80.

[3] Combs, T. and B. Bederson. Does Zooming Improve Image Browsing? *Proc. Digital Libraries 1999*, 130–137.

[4] Drucker, S., C. Wong, A. Roseway, S. Glenner, and S. De Mar. MediaBrowser: reclaiming the shoebox. *Proc. AVI 2004*, 433–436.

[5] Fertig, S., E. Freeman, and D. Gelernter. Lifestreams: An Alternative to the Desktop Metaphor. *Proc. SIGCHI 1996*, 410–411.

[6] Frohlich, D., A. Kuchinsky, C. Pering, A. Don, and S. Ariss. Requirements for Photoware. *Proc. CSCW 2002*, 166–175.

[7] Graham, A., H. Garcia-Molina, A. Paepcke, and T. Winograd. Time as essence for photo browsing through personal digital libraries. *Proc. Joint Conf. Digital Libraries*, 2002.

[8] Kang, H. and B. Shneiderman. Visualization methods for personal photo collections: Browsing and searching in the PhotoFinder. *Proc. IEEE Intl. Conf. on Multimedia and Expo*, 2000.

[9] Kuchinsky, A., C. Pering, M. Creech, D. Freeze, B. Serra, and J. Gwizdka. FotoFile: a consumer multimedia organization and retrieval system. *Proc. SIGCHI 1999*, 496–503.

[10] Mackinlay, J., G. Robertson, and R. DeLine. Developing calendar visualizers for the information visualizer. *Proc. UIST 1994*, 109–118.

[11] Perlin, K. and D. Fox. Pad: An Alternative Approach to the Computer Interface. *Proc. SIGGRAPH 1993*, 57–64.

[12] Plaisant, C., B. Milash, A. Rose, S. Widoff, and B. Shneiderman. LifeLines: Visualizing Personal Histories. *Proc. SIGCHI 1996*, 221–227.

[13] Platt, J. C., M. Czerwinski, and B. A. Field. PhotoTOC: Automatic Clustering for Browsing Personal Photographs. *Technical Report MSR-TR-2002-17*, Microsoft Research, 2002.

[14] Ringel, M., E. Cutrell, S. Dumais, and E. Horvitz. Milestones in Time: The Value of Landmarks in Retrieving Information from Personal Stores. *Proc. INTERACT 2003*.

[15] Rodden, K. and K. Wood. How Do People Manage Their Digital Photographs. *Proc. SIGCHI 2003*, 409–416.

[16] Suh, B., H. Ling, B. Bederson, and D. Jacobs. Automatic Thumbnail Cropping and Its Effectiveness. *Proc. UIST 2003*, 95–104.