

Lean and Zoom: Proximity-Aware User Interface and Content Magnification

Chris Harrison Anind K. Dey
 Human-Computer Interaction Institute
 Carnegie Mellon University
 Pittsburgh, PA 15213
 {charrison, anind}@cs.cmu.edu



Figure 1. On-screen content is magnified as a user leans closer to the screen.

ABSTRACT

The size and resolution of computer displays has increased dramatically, allowing more information than ever to be rendered on-screen. However, items can now be so small or screens so cluttered that users need to lean forward to properly examine them. This behavior may be detrimental to a user's posture and eyesight. Our Lean and Zoom system detects a user's proximity to the display using a camera and magnifies the on-screen content proportionally. This alleviates dramatic leaning and makes items more readable. Results from a user study indicate people find the technique natural and intuitive. Most participants found on-screen content easier to read, and believed the technique would improve both their performance and comfort.

Author Keywords

Lean, magnification, semantic zooming, user interfaces, assistive technology, proximity aware, posture, sensors.

ACM Classification Keywords

H5.2. [User Interfaces]: Interaction Styles, Graphical User Interfaces.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2008, April 5–10, 2008, Florence, Italy.
 Copyright 2008 ACM 978-1-60558-011-1/08/04...\$5.00.

INTRODUCTION

People naturally lean towards items in order to better inspect them. The magnification provided by being physically closer can expose otherwise unseen details. However, the cumulative effects of repeated leaning can be detrimental to users' posture [2]. Also, the need to focus on closer-than-usual surfaces can cause eyestrain, double vision, headaches and other vision-related problems [2]. This lean behavior is prevalent in computer use; display technology has advanced considerably, both in size and resolution, allowing more information to be rendered on screen than ever before. The magnitude of this problem is growing as people spend an ever-increasing amount of time in front of computers, both at work and home.

The most obvious way to counteract this problem would be to design user interfaces such that items are easily viewable and leaning is rendered unnecessary. However, this is impractical on several levels. Foremost, it is simply not possible to upgrade all software to have this desired behavior. Secondly, people have different thresholds for when leaning becomes necessary. An attempt to design a user interface that works for the lowest common denominator would likely result in a less productive interface for a majority of users. Lastly, the need for magnification is highly content- and user-specific. There is no reliable way to establish what content users would prefer to see with extra or less detail. For example, a photograph of a class reunion encountered online might warrant a close examination by some people, but not others.

Operating systems have provided zoom features for many years, although these are rarely accessed. Buttons to increase and decrease font size in web browsers have seen more success. However, graphical content and the user interface are not scaled, leaving some content small and often distorting the page layout. One possible reason for the limited use of these features is that they are not easily accessed or triggered. It may be easier for users to simply lean in and cope with the current appearance.

In order to sidestep these complexities, Lean and Zoom employs a different approach. In particular, the technique leverages users' existing lean habits while not requiring modification to the user interface.

INTERACTION TECHNIQUE

The Lean and Zoom system magnifies on-screen content in response to a user leaning forward (Figure 1). The zoom is directly proportional to the extent of the lean – the greater the lean, the greater the magnification. When zoomed in, the mouse is used to pan around the screen.

This interaction technique has several important qualities. Foremost, Lean and Zoom takes advantage of an unused input channel: lean. Second, the system has essentially no learning curve because the notion of leaning forward for visual enlargement is natural. Third, the technique reduces lean extent, by providing a greater zoom gain than the physical environment – people need to lean less in order to achieve an equivalent magnification level. Although not eliminating the need to lean, reducing the extent of leans may ease detrimental posture and eyesight effects [2]. Fourth, by relying entirely on users for activation, Lean and Zoom does not require any knowledge of on-screen content. The technique is only triggered when the user has deemed it necessary (or sufficiently interesting). Lastly, the graphical user interface does not need to be altered in anyway - magnification is achieved by capturing the current visual appearance and simply scaling it. This means the operating system, applications, and content need not be modified in any way for the technique to be applied.

SYSTEM DESCRIPTION

The current Lean and Zoom implementation relies on a camera to calculate a user's lean position. This is an attractive approach as computers with built-in or bundled cameras are becoming increasingly prevalent, allowing the technology to be readily and inexpensively deployed. A simple vision system tracks two points on the user's face or head; eyes are an obvious candidate. When the system is first started, the user's nominal posture is recorded by calculating the distance between the two points. As the user leans in, the distance between the two markers increases. This value is subtracted from the nominal posture distance and an appropriate zoom level is applied. Sensing resolution is dependent on the user's proximity and position. At typical viewing distances, a generic webcam captures a vertical plane approximately 50cm wide. Thus, a

camera with the horizontal resolution of 640 pixels allows horizontal movement as small as 0.8mm to be detected, providing a resolution in the lean dimension of around 4mm. Resolution increases as the user leans forward.

Lean and zoom uses a series of discrete zoom levels. This was a purposeful design choice. If user proximity is directly tied to a continuous zoom scale, the screen appears to pulsate, an effect caused by sensing inaccuracies and natural, subtle motions of the head. Moreover, using large, discrete zoom steps gives the user some freedom to move (e.g., shift position in chair) without the magnification level changing. A series of overlapping proximity windows was used to eliminate "flickering" near zoom level boundaries.

Figure 2 displays the sensed proximity of a user over approximately a four-minute period in blue (upper). The value is the horizontal distance in pixels between the two tracking points. Tracking errors, seen as oscillation and spikes, are visible. Various techniques are employed to compensate for sensing inaccuracies, for example, a moving average. The orange line (lower) shows the magnification level the system applied to the screen. The gain of the zoom can be adjusted as desired.

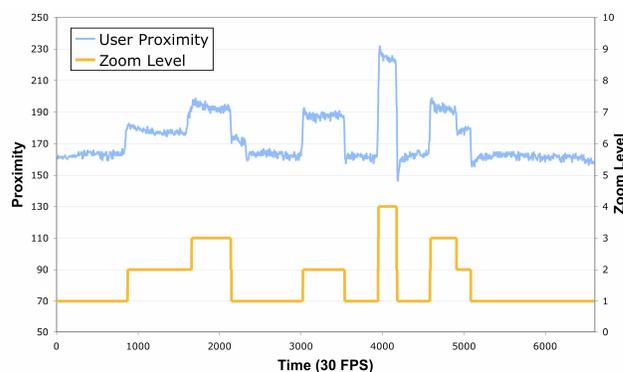


Figure 2. The sensed lean position of a user (top) and the zoom level applied to the screen (bottom).

USER STUDY

Ten participants were recruited using an email campaign and had a mean age of approximately 47 (min=24, max=70). No participants had any physical limitations and all used a computer on a regular basis. The experiment was completed on an Apple laptop with a 13" widescreen running at a resolution of 1280x800. The laptop included an integrated camera capable of capturing video at 30 frames per second at 640x480. An external two-button mouse with scroll wheel was provided.

Users were seated at a table with the laptop placed approximately 20cm from the edge. The laptop display was tilted such that it was perpendicular to the user's vision. This simultaneously oriented the camera. Participants were asked to wear a small plastic headband with two small colored markers. This allowed the vision system to track a user's proximity without needing to calibrate to the participant's eyes. A quick tutorial was provided and users were given a few minutes to experiment with the system.

Two model tasks were given to participants. Each task was completed twice, once with Lean and Zoom enabled and once with it disabled (*i.e.*, a regular user interface). One of the tasks required users to read and navigate the New York Times website for three minutes. This site was chosen for its text-rich nature. The task had no target objective for users to complete, and was primarily included to allow qualitative evaluation of the technique. Participants were encouraged to provide feedback about the system as they used it. The other task required users to point out and name four objects hidden in an illustration - discovery times were recorded. The task continued until the participant found all four items or three minutes had elapsed. The image and hidden objects changed when the task was repeated for the second time. In addition to exposing users to how the system would work with image-rich content, the task also provided a means to investigate if the Lean and Zoom technique had an impact on performance, most notably in object recognition and scanning speed. To compensate for order effects, the presentation order of these experimental conditions was balanced across participants.

Additionally, a simple semantic zooming application (requiring knowledge about content) was created for demonstration purposes (Figure 3). The application displayed instructions for assembling a workbench. As users leaned forward, additional details were rendered, such as part numbers and the placement of screws, bolts and other hardware. Other semantic zooming examples were explained verbally.

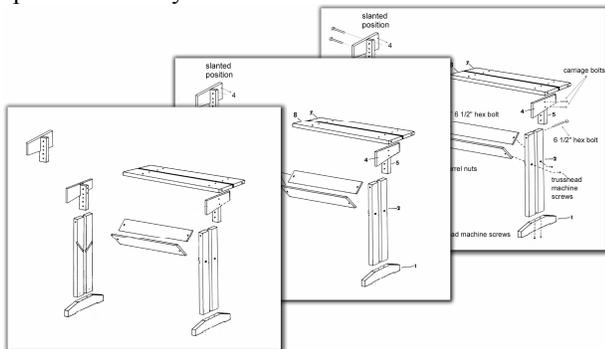


Figure 3. The three states of the demonstration semantic zooming application.

Participants were given two surveys, one before seeing the system and another at the conclusion of the study. A brief interview was conducted following the final survey to gather additional information about people's lean behavior and reaction to the interaction technique.

RESULTS AND DISCUSSION

Data from the first survey supported our initial assumption that people are universally and frequently leaning towards their computer displays. Participants estimated that they lean about once a day (20%), many times a day (70%), or many times an hour (10%). This estimation is based on what participants remember. However, since leaning occurs essentially subconsciously, the frequency may be much

greater in reality. Figure 4 illustrates how often users believed they lean for a variety of computer-based tasks. Participants also reported that viewing text and images were the primary causes for leaning in.

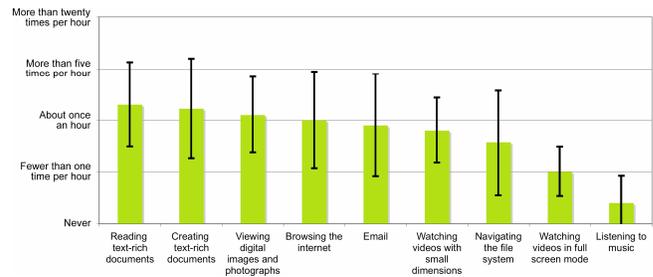


Figure 4. Participant-rated lean frequently for a variety of computer-based tasks.

Despite initial reservations, participants overwhelmingly favored the technique at the conclusion of the study. Many users commented that the system was “weird at first ... [but] like it now” and acknowledged there was a “newness problem,” but that they “could adjust very easily.” One participant exclaimed, “used to bigger text [now] – harder to go back.” Another participant, when reading the New York Times page without Lean and Zoom, commented, “[I] can’t believe how much I zoomed in [before].” The consensus was that items were “much easier to read” and generally “easier on your eyes.” Three participants who usually wore reading glasses or bifocals said the technique would allow them to use the computer without them.

Participants felt the interaction technique was useful in two distinct ways. The first was for “finding very small things” and examining items with “fine detail,” such as data-rich spreadsheets or maps. The technique was also useful for zooming in and “getting rid of extraneous information,” allowing them to look “[only] at the most important stuff.”

The concluding survey mirrored the users’ positive comments. On a scale of one being not at all intuitive and five being very intuitive, the mean response was 4.0 ($sd=1.2$). Using the same scale, navigating with the mouse when zoomed-in scored a 4.2 ($sd=1.1$). 80% of participants felt that on-screen content was easier to read when Lean and Zoom was available for use. Correspondingly, 70% of participants responded that they would be “very likely” to use the technique if it were a feature on their computer. Figure 5 provides a breakdown of how useful participants believed the technique would be for a variety of tasks.

Discovery times from the image-searching experiment did not reveal any significant performance effects. However, there are some indications that Lean and Zoom may marginally degrade performance. For example, participants found 8% fewer items when Lean and Zoom was available (3.1 vs. 3.4). Also, the average time for participants who located all four items was extended by 2% when Lean and Zoom was on (70.3 vs. 72.0 seconds). However, it is possible these effects are due to inexperience with the technique. Results from the concluding survey were very

favorable, with 90% of participants indicating that they believed the technique would increase their performance.

The concluding survey also asked users to consider which device categories they believed the technique would be applicable to. Every participant believed Lean and Zoom would be appropriate for desktop machines with detached monitors. Additionally, 90% of participants selected laptops and 80% mobile devices (e.g., cell phone, PDA). Another question allowed users to select which methods they believed would be most intuitive to enable and disable the interaction technique. The most popular choices were a special key that acted as a toggle (70%) and a physical gesture such as a head nod (30%). Surprisingly, 20% believed having no toggle method would be intuitive, where the feature was simply always active.

Participants were uniformly enthusiastic about the lean-augmented semantic zooming application. Every participant believed the technique and application combination would be useful, with more than half rating it as “very useful.”

Increasing user comfort was a chief goal of Lean and Zoom. Our findings suggest the system was successful in this regard; 80% of participants believed the technique would improve their comfort. However, 60% of participants thought having the technique available would cause them to lean more. This is seemingly contradictory to results from the initial survey, where a majority of participants believed that leaning forward was detrimental to both posture and eyesight (90% and 50% respectively). One explanation for this apparent paradox is that the content magnification provided by the technique causes people to lean less dramatically. For example, a user may only have to lean in 5° when using Lean and Zoom in order to achieve the same visual enlargement as an un-augmented 15° lean. Previous research has shown that negative posture-related effects intensify as lean angle increases [2].

The study also revealed that lean position is very stable. Once a person assumes a pose, they will hold that position until an adjustment is needed. During this time, the head moves very little. When people do move, the motion is fluid, allowing for more accurate tracking. Additionally, it is uncommon for someone to lean accidentally, reducing the possibility of false positives. Figure 2, discussed earlier, illustrates this stability. Furthermore, only a single participant believed their leaning was a natural, involuntary motion unrelated to screen content. Thus, it appears that lean behavior is generally associated with changes in content. These qualities make lean position a strong and natural candidate for input.

RELATED WORK

While there has been much work in supporting magnification in user interfaces, and in using computer vision to support user interaction, there has been little work in the intersection. Current applications often provide on-screen widgets that control the zoom level. Research efforts

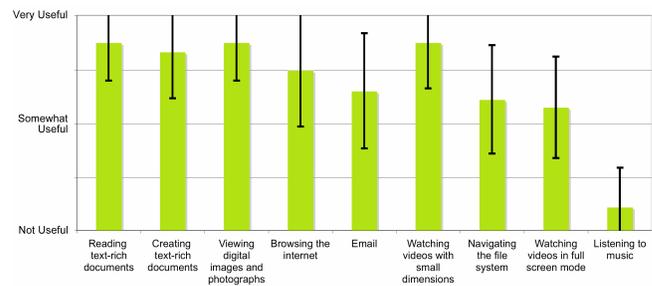


Figure 5. Participant-rated usefulness of Lean and Zoom for a variety of computer-based tasks.

have yielded zooming techniques that leverage the speed of scrolling [3] and the pitch and volume of a user’s voice [4]. Although our prototype system uses a linear, full-screen zoom, other zoom methods are applicable (e.g., hyperbolic zoom). Researchers have been applying computer vision for many years to support user interaction, for example, determining dwell time in front of an interface [6], tracking eye gaze [5], and capturing user gestures [1]. Additionally, tracking user motion and position has been popular in an entertainment context, where such data can be used to make games more engaging (e.g., Sony EyeToy™, [7]).

CONCLUSION AND FUTURE WORK

The Lean and Zoom system detects a user’s lean position and proportionally magnifies on-screen content. Results from a study indicate that users believe this interaction technique is intuitive, increases comfort, and improves performance. Future work includes investigating other proximity sensing technologies, examining the technique’s potential for people with limited mobility, and exploring the broader implications of proximity-aware user interfaces.

REFERENCES

1. Bobick, A. *et al.* The KidsRoom: A perceptually-based interactive and immersive story environment. *Presence* (8)4, 367-391.
2. Grandjean, E. *Fitting the Task to the Man*. Taylor & Francis, New York, NY, 1988.
3. Igarashi, T. and Hinckley, K. Speed-dependent automatic zooming for browsing large documents. *Proceedings of UIST 2000*, 139-148.
4. Igarashi, T. and Hughes, J.F. Voice as sound: Using non-verbal voice input for interactive control. *Proceedings of UIST 2001*, 155-156.
5. Jacob, R. What you look at is what you get: Eye movement-based interaction techniques. *Proceedings of CHI '90*, 11-18.
6. Vogel, D. and Balakrishnan, R. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. *Proceedings of UIST 2004*, 137-146.
7. Wang, S. *et al.* Face-tracking as an augmented input in video games: enhancing presence, role-playing and control. *Proceedings of CHI 2006*, 1097-1106.