Gaze-enhanced Scrolling Techniques

Manu Kumar, Terry Winograd
Stanford University, HCI Group
Gates Building, Room 382
353 Serra Mall
Stanford, CA 94305, USA
{sneaker, winograd}@stanford.edu

ABSTRACT
Scrolling is an inherent part of our everyday computing experience. Contemporary scrolling techniques rely on the explicit initiation of scrolling by the user. However, the act of scrolling is tightly coupled with the user's ability to absorb information via the visual channel. The use of eye gaze information is therefore a natural choice for enhancing scrolling techniques. We present several gaze-enhanced scrolling techniques for manual and automatic scrolling which use gaze information as a primary input or as an augmented input. We also introduce the use off-screen gaze-actuated buttons for document navigation and control.

ACM Classification: H5.2. User Interfaces. – Input devices and strategies. H5.2. User Interfaces: Windowing Systems, H5.m. Information Interfaces and presentation (e.g., HCI): Miscellaneous.

General terms: Human Factors, Algorithms, Performance, Design

Keywords: Scrolling, Automatic Scrolling, Eye Tracking, Gaze-enhanced Scrolling, Page Down, Gaze-enhanced Page Down, Off-screen Targets.

INTRODUCTION
Scrolling is an inherent part of our everyday computing experience. It is essential for viewing information on electronic displays, which provide a limited viewport to a virtually unlimited amount of information. Contemporary scrolling techniques rely on the explicit initiation of scrolling by the user. Considerable prior work [3, 4, 6, 10, 11] has been done in evaluating various techniques and devices for scrolling.

The act of scrolling is tightly coupled with the user's ability to absorb information via the visual channel, i.e. the user initiates a scrolling action to inform the system that he/she is now ready for additional information to be brought into view. We therefore posit that gaze information can be an invaluable source of contextual information making it a natural choice for enhancing scrolling techniques.

The eyes are a perceptual organ designed for the purpose of looking. Therefore, when using gaze information for input, it becomes necessary to disambiguate when the subject is simply looking around from when he or she intends to perform an action [5]; otherwise we end up with the Midas Touch problem: everywhere you look another command is activated.

By understanding the characteristics of reading patterns and how users consume visual information [2, 7, 8] it is possible to devise new techniques for scrolling, which can either use gaze-information to automatically control the onset and the speed of scrolling or use gaze information passively to augment manual scrolling techniques.

Individual differences in reading and scanning patterns may result in no one technique being suitable for all users. We therefore explore a range of techniques in this paper that may satisfy different user preferences. We designed and implemented techniques for both manual and automatic scrolling on a Tobii 1750 [9] eye tracker. We also introduce the use of off-screen gaze-actuated buttons or hotspots that allow users to explicitly control document navigation.

MANUAL SCROLLING
The Page Up / Page Down Problem
The implementation of Page Up and Page Down on contemporary systems is based on the expectation that the user will press the page down key when he or she is looking at the last line on the page. However, our contextual inquiry revealed that users often initiate scrolling in anticipation of getting towards the end of the content in the viewport. This results in users pressing page down before reaching the last line of the text. Consequently, the text the user was looking at scrolls out of view off the top of the viewport. This necessitates a fine-tuning of the scrolling movement to bring the text back into view. In addition, most users tend to lose track of where they were reading once the page scrolls and must reacquire their position in the text.

Gaze-enhanced Page Up / Page Down
We propose a new approach for a gaze-enhanced page-down which uses a *GazeMarker* to always keep users' eyes on the text they were reading even through page transitions.

In this approach, the user's eye gaze on the screen is tracked. When the user presses the page down key, the region where the user was looking immediately before pressing the page down key is highlighted. We call this highlight a "GazeMarker". The page is then scrolled such that the highlighted region becomes the topmost text shown in the
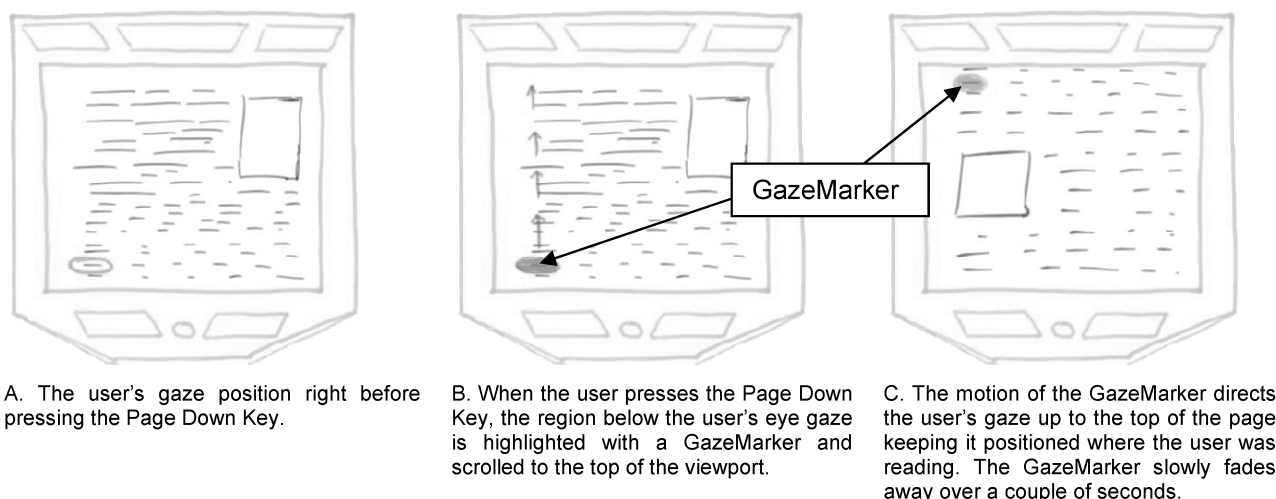
A. The user's gaze position right before pressing the Page Down Key.

B. When the user presses the Page Down Key, the region below the user's eye gaze is highlighted with a GazeMarker and scrolled to the top of the viewport.

C. The motion of the GazeMarker directs the user's gaze up to the top of the page keeping it positioned where the user was reading. The GazeMarker slowly fades away over a couple of seconds.

Figure 1. The Gaze-enhanced Page Up / Page Down approach addresses the limitations of current Page Up and Page Down Techniques by Positioning the region under the user's gaze at the bottom or top of the page respectively.

viewport (Figure 1). Since the highlight appears immediately before the page scrolls and then moves up in the viewport, the user's gaze naturally follows the highlight. This ensures that the user's gaze is kept on the text he or she was reading and minimizes the need to reacquire the text after scrolling. The GazeMarker slowly fades away within a few seconds.

AUTOMATIC SCROLLING

The design of any automatic scrolling techniques must overcome two main issues: a) the Midas Touch problem, b) controlling the speed at which the content is scrolled. We address each of these problems below.

Explicit Activation/Deactivation

PC keyboards include a vestigial Scroll Lock Key which the vast majority of users have never used. To overcome the Midas Touch problem we chose to use explicit activation of any of the automatic scrolling techniques by putting the Scroll Lock key back into use. The user toggles the automatic scrolling on and off by pressing the Scroll Lock key on the keyboard.

Estimation of Reading Speed

For several of the techniques presented, it is useful to be able to measure the user's reading speed. Previous work [2, 8] has shown that the typical eye movements (fixations and saccades) for a subject reading text conforms to Figure 2. Beymer et al. [2] present an estimate of reading speed based on forward-reads. For our use – to control scrolling -
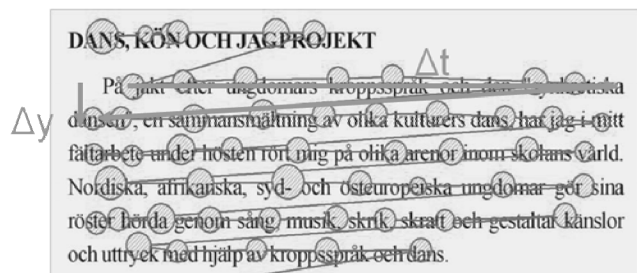

Figure 2. Estimation of reading speed. Vertical pixels viewed per second = $\Delta y/\Delta t$.

it is more interesting to measure the speed at which the user is viewing vertical pixels. This can be estimated by measuring the amount of time for the horizontal sweep of the user's eye gaze (Δt) and the delta in the number of vertical pixels during that time(Δy). The delta in the vertical pixels divided by the amount of time for the horizontal sweep ($\Delta y/\Delta t$) provides an instantaneous measure of "reading speed". A smoothing algorithm is applied to the instantaneous reading speed to account for variations in column sizes and the presence of images on the screen. The resulting smoothed reading speed provides a best guess estimate of the rate at which the user is viewing information on the screen.

We present three scrolling techniques that start and stop scrolling automatically, depending upon the user's gaze position. The techniques differ in the details of whether the content is scrolled smoothly or discretely. Each of the techniques presented scrolls text only in one direction. This was a conscious design choice to overcome the Midas Touch problem. Scrolling backwards or navigating to a particular section of the document can be achieved either by using manual methods or by using off-screen navigation buttons.

Eye-in-the-middle

The eye-in-the middle technique for automatic scrolling measures the user's reading speed while dynamically adjusting the rate of the scrolling to keep the user's gaze in the middle third of the screen (Figure 3). This technique relies on accelerating or decelerating the scrolling rates to match the user's instantaneous reading speed. It is best suited for reading text-only content since the user's scanning patterns for images included with the text may vary. This technique requires that the user read text while it is scrolling smoothly.

Smooth scrolling with gaze-repositioning

This automatic scrolling approach relies on using multiple invisible threshold lines on the screen (Figure 4). When the user's gaze falls below a *start threshold*, the document begins to scroll slowly. The scrolling speed is set to be

Figure 3. The eye-in-the-middle automatic scrolling technique adjusts the scrolling speed to match the user's reading speed and tries to keep the user's eyes in the middle thrid of the screen.
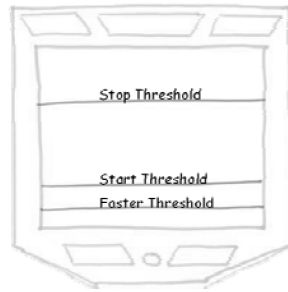

Figure 4. The smooth scrolling with gaze-repositioning technique allows for reading and scanning of content. Scrolling starts and stops depending on the position of the user's gaze with respect to invisible threshold lines on the screen.

slightly faster than the user's reading speed so as to gradually move the user's gaze position towards the top of the screen. When the user's gaze reaches a *stop threshold*, scrolling is stopped (text is stationary) and the user can continue reading down the page normally. If the user's gaze falls below a *faster threshold*, the system begins to scroll the test more rapidly. The assumption here is that either the scrolling speed is too slow or the user is scanning and therefore would prefer that the content scroll faster. Once the user's gaze rises above the start threshold, the scrolling speed is reduced to the normal scrolling speed. The scrolling speed can be adjusted based on each individual's reading speed.

In our implementation, the position of the threshold lines was determined based on user-feedback. In particular, placing the stop threshold line higher on the screen resulted in users worrying that the text would run away before they would have the chance to finish reading it. We therefore lowered the stop threshold to one-third the height of the screen so that scrolling would stop before the users became anxious. In addition, whenever scrolling is started or stopped, it is done by slowly increasing or decreasing the

scrolling rate respectively. This is done to make the transitions from one state to another continuous and fluid.

This approach allows for both reading and scanning, however, in this approach while the user is reading, sometimes the text is moving and other times the text is stationary.

Discrete scrolling with gaze-repositioning
The discrete scrolling with gaze-repositioning approach leverages the gaze-enhanced Page Up / Page Down technique for manual scrolling and extends it by adding an invisible threshold line towards the bottom of the screen. When the user's eyes fall below the threshold the system issues a page down command which results in the Gaze-Marker being drawn and the page being scrolled (Figure 5). The user's gaze must stay below the threshold for a micro-dwell duration (~150-200ms) before the event triggers. This minimizes the number of false activations from just looking around at the page and disambiguates scanning the screen from reaching the end of the content on the screen while reading. The scrolling motion happens smoothly to keep the user's eyes on the GazeMarker, but fast enough for the scrolling to appear as if it occurred a page at a time.

This approach ensures that users read only when the content is stationary (in contrast to the previous automatic scrolling approaches).

OFF-SCREEN GAZE-ACTUATED BUTTONS
The Tobii eye-tracker provides sufficient field of view and resolution in order to be able to clearly identify when the user is looking beyond the edges of the screen at the bezel. This provides ample room to create gaze-based hotspots to include various navigation controls. We implemented several variations of off-screen gaze-actuated buttons for document navigation as seen in Figure 6.

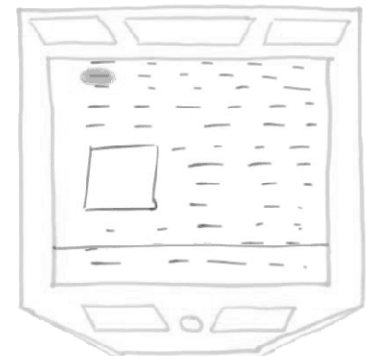### Dwell vs. Micro-Dwell based activation
Document navigation actions have either a discrete nature (such as the Home, End, Page Up and Page Down buttons), which require one time activation or have a more continuous nature such as the cursor keys on the keyboard or the controls on a scroll bar. The latter typically require the action to be performed repeatedly.

A. The user's gaze position when the eye gaze drops below the scrolling threshold.
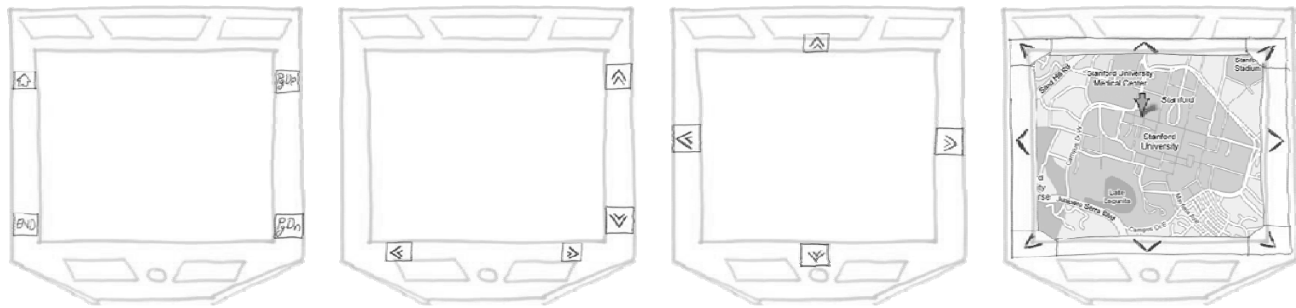

B. If the user's gaze stay below the threshold for the duration of a micro-dwell (~150-200ms) the system issues a Page Down command, which results in the GazeMarker being drawn.

C. The motion of the GazeMarker directs the user's gaze up to the top of the page keeping it positioned where the user was reading. The GazeMarker slowly fades away over a couple of seconds.

Figure 5. The discrete scrolling with gaze-repositioning leverages the gaze-enhanced Page Up / Page down and triggers a Page Down event when the users gaze falls below a threshold line for a specified duration.

A. Home, End, Page Up and Page Down buttons activated by dwell (400-500 ms)	B. Scrolling buttons activated by a micro-dwell (150-200ms) provide continuous input while the user is looking at them.	C. Scrolling buttons located in the center to be aligned with the user's gaze direction.	D. 8-way panning regions activated by looking slightly off-screen for a micro-dwell duration (150-200ms)

Figure 6. Off-screen gaze-actuated buttons/hotspots for document navigation and control. Buttons which trigger discrete events (Home, Page Down etc) use a dwell-based activation. Hotspots that have a more continuous action (scroll up etc) use a micro-dwell based activation.

In order to accommodate the different forms of these actions we implement two different activation techniques. The first, dwell-based activation, triggers only once when the user has been staring at the target for at least 400-500ms. For the actions which require continuous input, we chose to use a micro-dwell based activation where the user has been staring at the target for at least 150-200ms.

EVALUATION

Pilot studies with 10 users indicated that all subjects preferred the gaze-enhanced Page Up/Page Down technique over the normal Page Up / Page Down.

In studies with the automatic scrolling techniques, subjects felt that scrolling started when they expected it to, that the scrolling speed was neither too fast nor too slow and they felt that they were in control. Some subjects commented that they initially found it disconcerting to read while the text was moving, but that they got used to it with practice. It is conceivable that, like a teleprompter, subjects may be comfortable with reading moving text with practice. The discrete scrolling with gaze-repositioning technique overcomes the issues of reading moving text and was therefore he preferred approach for gaze-enhanced scrolling. Subjects indicated that they would prefer the gaze-enhanced scrolling techniques over traditional scrolling methods.

Systematic studies will need to be done to rigorously evaluate the benefits and shortcomings of these techniques.

CONCLUSION

We presented several techniques for gaze-enhanced scrolling which include augmenting existing manual scrolling techniques with gaze input and variations of automatic gaze-based scrolling techniques. We also introduced the use of off-screen gaze-actuated buttons or hotspots.

Gaze enhanced scrolling has the potential to radically reduce the number of scrolling actions users need to perform in order to surf the web or consume other information displayed in electronic form. With the inclusion of cameras into current display devices [1] and the impending reduction in cost of eye-tracking technology, gaze-based scrolling techniques will increase in importance and provide users with a natural alternative to current approaches.

REFERENCES
1. *Apple MacBook iSight camera.* Apple Computer: Cupertino, California, USA.
    http://www.apple.com/macbook/isight.html
2. Beymer, D. and D. M. Russell. WebGazeAnalyzer: A System for Capturing and Analyzing Web Reading Behavior Using Eye Gaze. In Proceedings of *CHI*. Portland, Oregon, USA: ACM Press. pp. 1913-16, 2005.
3. Cockburn, A., J. Savage, and A. Wallace. Tuning and Testing Scrolling Interfaces that Automatically Zoom. In Proceedings of *CHI*. Portland, Oregon, USA: ACM Press. pp. 71-80, 2005.
4. Hinckley, K., E. Cutrell, S. Bathiche, and T. Muss. Quantitative analysis of scrolling techniques. In Proceedings of *CHI*. Minneapolis, Minnesota, USA: ACM Press. pp. 65-72, 2002.
5. Jacob, R. J. K. The Use of Eye Movements in Human-Computer Interaction Techniques: What You Look At is What You Get. In Proceedings of *ACM Transactions in Information Systems*. pp. 152-69, 1991.
6. Laarni, J. Searching for Optimal Methods of Presenting Dynamic Text on Different Types of Screens. In Proceedings of *NordiCHI*. Arhus, Denmark: ACM Press. pp. 219-22, 2002.
7. Poynter Institute and Eyetools, Inc., *Eyetrack III: Online News Consumer Behavior in the Age of Multimedia*, 2004. http://poynterextra.org/eyetrack2004/index.htm
8. Rayner, K. Eye Movments in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin* 124(3). pp. 372-422, 1998.
9. Tobii Technology, AB, *Tobii 1750 Eye Tracker*, 2006. Sweden. http://www.tobii.com
10. Wallace, A., J. Savage, and A. Cockburn. Rapid Visual Flow: How Fast Is Too Fast? In Proceedings of *5th AUIC: Australasian User Interface Conference*. Dunedin: Australian Computer Society, Inc. pp. 117-22, 2004.
11. Zhai, S., B. A. Smith, and T. Selker. Improving Browsing Performance: A study of four input devices for scrolling and pointing tasks. In Proceedings of *IFIP Interact*. Sydney, Australia. pp. 286-92, 1997.