
Gaze-enhanced Scrolling Techniques

Manu Kumar

Stanford University, HCI Group
Gates Building, Room 382
353 Serra Mall
Stanford, CA 94305-9035
sneaker@cs.stanford.edu

Andreas Paepcke

Stanford University, HCI Group
Gates Building, Room 426
353 Serra Mall
Stanford, CA 94305-9035
paepcke@cs.stanford.edu

Terry Winograd

Stanford University, HCI Group
Gates Building, Room 388
353 Serra Mall
Stanford, CA 94305-9035
winograd@cs.stanford.edu

Abstract

We present several gaze-enhanced scrolling techniques developed as part of continuing work in the GUIDe (Gaze-enhanced User Interface Design) project. This effort explores how gaze information can be effectively used as input that augments keyboard and mouse. The techniques presented below use gaze both as a primary input and as an augmented input in order to enhance scrolling and panning techniques. We also introduce the use of off-screen gaze-actuated buttons which can be used for document navigation and control.

Keywords

Scrolling, Automatic Scrolling, Panning, Automatic Panning, Eye Tracking, Gaze-enhanced Scrolling, Gaze-enhanced Panning.

ACM Classification Keywords

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

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.

Copyright is held by the author/owner(s).

CHI 2007, April 28–May 3, 2007, San Jose, California, USA.

ACM 978-1-59593-642-4/07/0004.

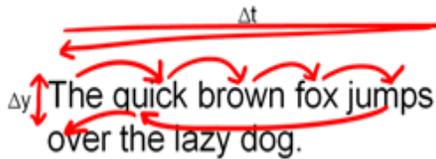


Figure 1. The amount of time for a round trip sweep along the X-axis defines Δt and the change in vertical pixels defines Δy . The instantaneous reading speed is measured as $\Delta y/\Delta t$

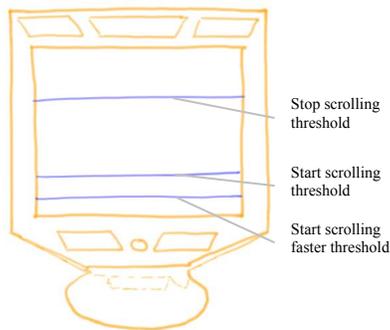


Figure 2. Thresholds for *smooth scrolling with gaze repositioning*. Scrolling starts when the user's gaze drops below the start threshold and stops when the user's gaze is above the stop threshold. Smooth continuous speeding allows users to read while the text is moving.

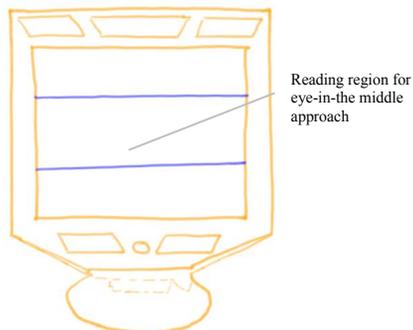


Figure 3. The eye-in-the middle approach dynamically adjusts the scrolling speed so as to always maintain the user's eye gaze within the reading region (except at the beginning and end of the document).

Considerable prior work [4-6, 12, 13] has been done on evaluating various techniques and devices for scrolling.

Our work is a direct outcome of our observation that 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 eye-gaze information can be an invaluable source of information for enhancing scrolling techniques.

By understanding the characteristics in reading patterns and how users consume visual information [3, 9, 10], it is possible to devise new techniques for scrolling, which can use gaze-information to automatically control the onset and speed of scrolling or use gaze information passively to enhance manual scrolling techniques.

In particular, we present several techniques that begin and end scrolling automatically, depending on the user's gaze position. The techniques differ in the details of whether the content is scrolled smoothly or discretely. In the case of smooth scrolling, the scrolling speed is adjusted based on the user's estimated reading speed, allowing the user to read the text while it is in motion. In the case of discrete scrolling, the user only reads while the content is stationary. We also explore virtual, gaze-actuated buttons that allow users to explicitly initiate scrolling. Comfort and subjective preference are the main criteria for evaluation of the techniques presented. We expect that individual differences in reading and scanning patterns will result in no one technique being suitable for all users. We therefore are experimenting with a range of techniques that may satisfy different user preferences.

Estimating Reading Speed

For several of the techniques presented, it is useful to be able to measure the user's reading speed. Beymer et al. [3] have showed that the fixation pattern for reading for a majority of users conforms to Figure 1. They also present an estimate of reading speed based on forward-reads. For our use – to control scrolling – 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.

Gaze-enhanced Scrolling Techniques

We present four gaze-informed scrolling techniques that we prototyped and tested:

Smooth scrolling with gaze-repositioning

This approach relies on using multiple invisible threshold lines on the screen (Figure 2). When the user's gaze falls below a start-scrolling threshold, the document begins to scroll slowly. The scroll speed is set such that the user is able to comfortably read the text even as it scrolls. The scrolling speed is set slightly faster than the user's reading speed so as to slowly move the user's gaze position towards the top of the screen. When the user's gaze reaches a stop-scrolling threshold, scrolling is stopped (text is stationary) and

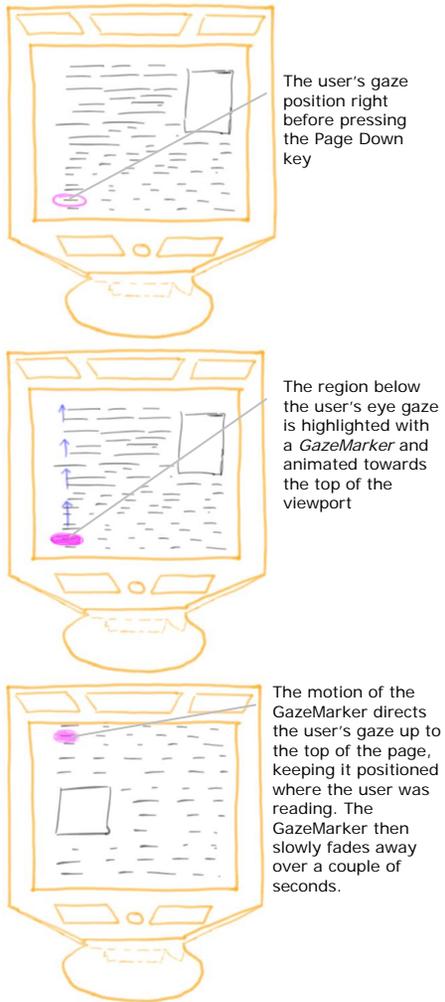


Figure 4. The *Gaze-enhanced Page Up / Page Down* approach addresses the limitations of current Page Up and Page Down techniques, by positioning the region under the user's gaze at the bottom or top of the page respectively.

The *Discrete scrolling with gaze-repositioning* approach leverages this technique to automatically issue a page down command when the user's is reading below a low threshold on the screen.

the user can continue reading down the page normally. If the user's gaze falls below a fast-scrolling threshold, the system begins to scroll the text more rapidly. The assumption here is that either the slow scrolling speed has been too slow or the user is scanning and therefore would prefer the faster scrolling speed. Once the user's gaze rises above the start-scrolling threshold, the scrolling speed is reduced to the normal scrolling speed. The scrolling speed can be adjusted based on each individual's reading speed.

Eye-in-the-middle

This approach measures the user's reading speed and dynamically adjusts the rate of the scrolling to keep the user's gaze in the middle third of the screen while reading. This technique relies on accelerating or decelerating the scrolling rates based on the instantaneous reading speed. It is best suited for reading text-only content since user's scanning patterns for images included with text may vary.

Gaze-enhanced Page Up / Page Down

The implementation of Page Up/Page Down on contemporary systems is based on the expectation that the user will only press the page down key when he or she is looking at the last line on the page. However, we have found that users often initiate scrolling by pressing page down in anticipation of getting towards the end of the content in the viewport. This often results in the user pressing page down before reaching the last line of the text. Consequently, the text the user was looking at is scrolled off the top of the viewport. The user then has to fine tune the scrolling (for example by using the arrow keys) to bring the text back into view and then reacquire where he/she left off reading.

We propose a new approach for a gaze-enhanced page-down: The user's 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 viewport. 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, minimizing the need to reacquire the text after scrolling. The *GazeMarker* slowly fades away within a few seconds.

Discrete scrolling with gaze-repositioning

This approach leverages the gaze-enhanced Page Up/Page Down approach. When the user's eyes fall below a pre-defined threshold on the screen, the system sends a page down command which results in the *GazeMarker* being drawn and the page being scrolled. 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 are only reading the text when it is stationary (as opposed to the smooth scrolling or eyes-in-the-middle approaches described before).

Use of Off-screen targets

An 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 three variations on this theme:

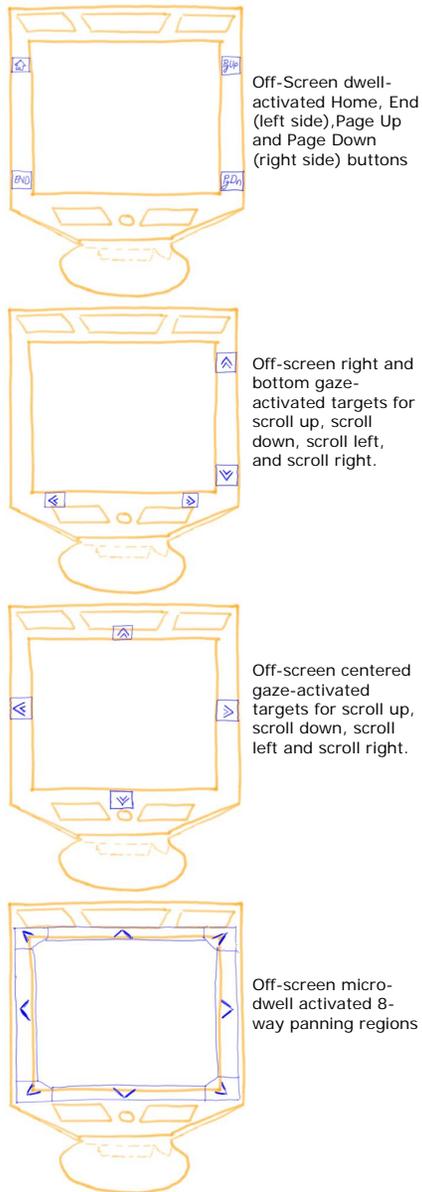


Figure 5. Off-screen targets for gaze based navigation, scrolling and panning.

Off-screen buttons for navigation

This technique provides off-screen buttons for Home and End on the left bezel of the screen and Page Up and Page Down on the right bezel of the screen. The off-screen buttons use dwell-based activation with an audio beep providing feedback when activated. The dwell time is set to 450ms as determined by previous research [7, 8]. Off-screen target navigation buttons work reliably, and have minimal issues with false activations due to their location and the use of dwell-based activation. The benefit over using a key on the keyboard is minimal, but it is a welcome step towards hands-free reading.

Off-screen Right/Bottom Scrolling Targets

Provides up, down, left and right scroll bar buttons, which are located off-screen along the right and bottom of the screen. The location of the targets is therefore similar to the location of scrollbars on the screen. The buttons are duration activated (i.e. not dwell). The up, down, left and right arrow keys are sent for as long as the user's gaze is on the button. The user's peripheral vision allows the user to tell if the screen contents have moved far enough.

Off-screen Center Scrolling Targets

Provides up, down, left and right scroll bar buttons, which are located off screen but in the center of the screen on all four sides. As opposed to the previous technique, in this approach the location of the targets is in a position where the user would look when planning to scroll in that direction. As before, these buttons are duration activated (i.e. not dwell). It should be noted that the entire bezel can be used as a target for activation of the scrolling as will be seen in the following approach.

Off-screen 8-way Panning Regions

As an extension to the off-screen center scrolling targets we also implemented 8-way panning by defining micro-dwell activated panning regions. This approach still uses dwell based activation to prevent false triggering when the user is looking around, but the duration of the dwell is reduced (~200 ms as opposed to 450ms) in order to make the panning begin as soon as the user fixates on the panning region. The panning region is defined to include a small sliver of the active screen region and the bezel of the screen.

We have used this technique to provide a virtual screen size that is larger than what can be accommodated on the current display. When the user looks at the trigger region for a duration longer than the micro-dwell duration, the screen pans to reveal the off-screen regions. We have also extended this approach to using virtual desktops, where the user can bring a virtual desktop into view simply by looking off-screen in the direction of the virtual desktop location.

Implementation

We implemented prototypes for each of the above mentioned techniques on a Tobii 1750 eye gaze tracker [11]. Most implementations provide a general solution that can be used across all applications. The gaze-enhanced Page Up / Page Down implementations were implemented as an extension to Mozilla Firefox using Greasemonkey [2] scripts.

Preliminary Evaluation

We have conducted pilot studies to gauge user reaction to the gaze-enhanced scrolling techniques described above. Feedback from the pilot studies was used to help refine the techniques and motivated key design

changes (such as the introduction of micro-dwell). Detailed evaluation of the scrolling techniques is planned as future work. We conducted a usability evaluation of the smooth scrolling with gaze-repositioning technique.

Method

To evaluate the smooth scrolling with gaze-repositioning technique we conducted a two part study with 10 subjects (6 male, 4 female). The average age of the subjects was 22 years. None of the subjects wore eye-glasses, though 2 did use contact lenses. None of the subjects were colorblind. English was the first language for all but two of the subjects. On average, subjects reported that they did two-thirds of all reading on a computer. The scroll-wheel was the most-favored technique for scrolling documents when reading online, followed by scroll bar, spacebar, page up / page down or arrow keys.

In the first part of the study, subjects were told that they would be trying a new gaze-based automatic scrolling technique to read a web page. For this part of the study, subjects were given no explanation on how the system worked. To ensure that subjects read each word of the document, we requested them to read aloud. We did not test for comprehension of the reading material since we were only interested in the subjects being able to view the information on the screen. Once subjects had finished reading the page, they were asked to respond to questions on a 7-point Likert scale.

In the second part of the study, we explained the technique's behavior to the subjects and showed them the approximate location of the invisible threshold lines (Figure 2). Subjects were allowed to practice and

become familiar with the approach and then asked to read one more web page. At the conclusion of this part subjects again responded to the same set of questions as before.

Results

Figure 6 summarizes the results from the study showing the subjects' responses in each of the two conditions (without explanation and with explanation).

Subjects' feeling that *scrolling started when they expected it to*, and that *they were in control* show increases in the with-explanation condition. For all other questions regarding comfort, fatigue and user preference there was no significant change in the subjects' responses across the two conditions. Subjects' response on the reading speed was mostly neutral, suggesting that they felt the scrolling speed was reasonable. While the differences in the results for reading speed in the two conditions are not significant, results do show that subjects were more comfortable (more neutral) in the with-explanation condition since they were more familiar with the operation and less worried about the content running off the screen.

Several subjects commented that they found reading text while it was scrolling to be disconcerting at first, but then became more comfortable with it once they realized that the text would not scroll off the screen and would stop in time for them to read. Based on this observation we hypothesize that subjects may prefer the discrete scrolling with gaze repositioning approach since it would require them to read only when the text was stationary.

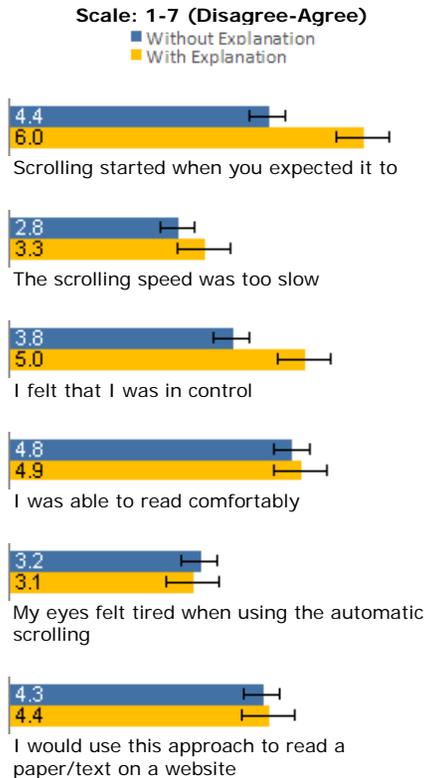


Figure 6. Subjective evaluation results for *Smooth scrolling with gaze-repositioning* in two conditions (with and without explanation of how the system works). Error bars show Standard Error.

Future Work

As part of the continuing research in this area, we will be conducting formal usability evaluations of the techniques described above. Evaluations will compare and contrast techniques to existing non-gaze-based scrolling techniques and also compare the gaze-enhanced techniques amongst themselves. Additional information regarding document structure (paragraph endings, location of text/images, size of the text, column structure etc.) may also be used to create content-sensitive gaze-enhanced scrolling techniques.

Conclusion

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. We hope that this paper will help to encourage further interest and research in developing gaze-based scrolling techniques.

Acknowledgements

The authors would like to thank Shumin Zhai and David Beymer for several insightful discussions on this topic.

References

1. *Apple MacBook iSight camera*. Apple Computer: Cupertino, California, USA.
<http://www.apple.com/macbook/isight.html>
2. *Greasemonkey Firefox extension*.
<http://greasemonkey.mozdev.org/>
3. 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.
4. 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.
5. 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.
6. Laarni, J. Searching for Optimal Methods of Presenting Dynamic Text on Different Types of Screens. In Proceedings of *NordiCHI*. Arhus, Denmark: ACM Press, 2002.
7. Majaranta, P., A. Aula, and K.-J. R  ih  . Effects of Feedback on Eye Typing with a Short Dwell Time. In Proceedings of *ETRA: Eye Tracking Research & Applications Symposium*. San Antonio, Texas, USA: ACM Press. pp. 139-46, 2004.
8. Majaranta, P., I. S. MacKenzie, A. Aula, and K.-J. R  ih  . Auditory and Visual Feedback During Eye Typing. In Proceedings of *CHI*. Ft. Lauderdale, Florida, USA: ACM Press. pp. 766-67, 2003.
9. Poynter Institute and Eyetools, Inc., *Eyetrack III: Online News Consumer Behavior in the Age of Multimedia*, 2004.
<http://poynterextra.org/eyetrack2004/index.htm>
10. Rayner, K. Eye Movments in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin* 124(3). pp. 372-422, 1998.
11. Tobii Technology, AB, *Tobii 1750 Eye Tracker*, 2006. Sweden. <http://www.tobii.com>
12. 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.
13. 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.