Gaze-Contingent Screen Magnification Control: A Preliminary Study

Roberto Manduchi[1] and Susana Chung[2]

[1] University of California, Santa Cruz, USA
[2] University of California, Berkeley, USA
manduchi@ucsc.edu

Abstract. People with low vision often use screen magnification software. Screen magnification requires continuous control of the onscreen content by moving the focus of magnification with the mouse or the trackpad. In this contribution, we explore the possibility of controlling the focus of magnification by means of the user's own eye gaze, which is measured by a commercial gaze tracker. We conducted two small experimental studies with individuals with impaired central vision, who used two screen magnification modalities to read two different types of documents. In the first study, mouse tracks and gaze point tracks were collected during manual control for later analysis. In the second study, the center of magnification was controlled by the user's own gaze, using two different control mechanisms. This preliminary study highlights the potentials and shortcomings of gaze-contingent screen magnification control for easier access of onscreen content with low vision.

Keywords: Screen magnification; Eye gaze tracking; Gaze-contingent display.

1 Introduction

Many people living with low vision use screen magnifiers to read documents and web pages on a computer. As more and more textual content is consumed online rather than in printed form, screen magnifiers are taking on the role of more traditional desktop video magnifiers (sometime called CCTV magnifiers), which have been used for decades to access printed text In many ways, a screen magnifier functions like a magnifying glass – while being purely software-based. It is the tool of choice for those individuals with some functional vision who do not need to (or choose not to) resort to screen readers. Multiple types of screen magnification are available on the market, either integrated in operating systems (Windows or MacOS), or in the form of specialized software such as ZoomText or MAGic.

Screen magnification is a powerful access technology, but it is not without its shortcomings. The common crux of screen magnification is that it requires continuous manual scrolling (using the mouse or trackpad) in order to move the focus of magnification (located at the mouse cursor), which determines the portion of the document to be magnified. Continuous scrolling of magnified content may represent a burden for the viewer (*page navigation problem* [4]). Manual scrolling often results in slow reading [8] and can be challenging for those who don't have full motor control of their hands.

The need for continuous manual scrolling of the magnified content could be mitigated by technology designed to assist the user in moving the focus of magnification. In this contribution, we propose a new system of gaze-based magnification control. The proposed system enables scrolling control by means of the viewer's own gaze, which is computed by an eye gaze tracker. This hands-free modality has the potential to afford a more natural experience when reading onscreen content than standard approaches that require use of mouse or trackpad.

This contribution describes two studies, comprising an initial data collection experiment, followed by a preliminary test of simple gaze-contingent magnification control algorithms. In Study 1, six participants with low vision operated two types of customized screen magnification software to read text from two onscreen documents. The application recorded the mouse tracks as well as the gaze point tracks, which were measured by an IR gaze tracking device. This study was meant to inform the design of a mechanism that uses gaze data to control the location of the focus of magnification. In Study 2, three simple mechanisms of gaze-contingent magnification control were tested by three participants with low vision. The goal of this preliminary study was to evaluate the feasibility of gaze-based magnification control, and to provide indications for future research in this direction.

2 Related Work

Prior research (e.g., [7,8]) studied the performance of onscreen reading for people with low vision (sequential reading as well as non-sequential skipping or skimming modalities [5]) using different types of magnification mechanisms, with outcomes expressed in terms of reading speed or error rates.

Gaze-contingent mechanisms for image enhancement/magnification are designed to process images at the location of the gaze point, or possibly at the preferred retinal locus of individuals with central field loss. Various image processing functions have been considered in the literature, including: "bubble" (or "fisheye" [3]) filters, which shift the image area hidden by a scotoma to a nearby peripheral area [1]; band-limited contrast enhancement [13]; adjustment of letter spacing to minimize "crowding" in peripheral areas [1]; and Region of Augmented Vision selection and magnification [2]. For the systems cited above to work, a high-precision gaze tracker (with resolution as high as 0.1°) is generally needed [14]. This normally requires head stabilization (e.g., by means of a chin rest) to ensure precise gaze tracking, or implementation in a head-mounted display. The need for head stabilization or head-mounted display greatly reduces the practical appeal of such devices. In contrast, we aim to build a system that is easy to use in a natural viewing setting, and that could benefit a variety of users with low vision, rather than only those with central field loss. By employing a relatively large screen lens or full screen magnification, rather than a highly localized "bubble", we afford the use of low-accuracy gaze trackers that do not require expensive hardware, and allow for some amount of head motion during reading. Experiments pairing a gaze tracker with magnification software were presented in [10-12]. These systems are similar to the Screen Lens–Integrative (SL–I) modality discussed in Sec. 4.1.

3 Study 1: Data Acquisition

3.1 Method

Participants. We recruited six participants with low vision for this study from the optometry clinic in our university. The participants had varying prior experience with screen magnification. P1, P4 and P6 were accustomed to magnifying the content of a Word document by "zooming" on the trackpad. P2 regularly used the AI Squared ZoomText Reader software (P4 also used this software on occasion). P3 mentioned that she normally increased the font size of a document (Ctrl + on Mac) for better reading. P5, who rarely used a computer, never used screen magnification before. Two participants (P1 and P2) used eyeglasses during the experiment.

Apparatus. We created a screen magnification software application for Windows 10, using the Magnification API and the Tobii EyeX Engine. The application ran on a Dell Latitude 3470 laptop computer, with screen size of 31 x 17.5 cm, and resolution of 1366 x 768 pixels. The computer was connected to a Tobii X2-30 eye tracker, attached to the lower edge of the screen. The X2-30 tracker captures data at 30 Hz and has nominal accuracy of 0.4˚ in ideal situations. It does not require head stabilization, which makes it suitable for "real world" applications. For the eye tracker to function correctly, a per-user prior calibration phase is necessary. This operation, which may take a few minutes, requires the user to follow with their gaze a dot moving on the screen, until prompted by the system that calibration has been completed.

Our application allows one to select between full screen (FS) and screen lens (SL) (sometimes called picture-in-picture) magnification. Full screen magnification expands the content of the whole screen around the focus of magnification (FoM), which coincides with the location of the mouse cursor. This results in only a portion of the onscreen content being visible within the screen viewport. Screen lens magnification uses the paradigm of a magnifying glass to only enlarge a rectangular portion of the screen (note that also in this case, the FoM coincides with the location of the cursor as controlled by the mouse). In both modalities, participants were able to select the desired magnification factor (over a logarithmic scale) using the keyboard. For the screen lens modality, participants were able to vary the width and the height of the rectangular "lens", still using the keyboard. The application captured all mouse movements as well as all measured gaze points (i.e., the points on the screen where gaze was directed, as estimated by the gaze tracker).

Experiments. Each participant underwent a sequence of four trials. In each trial, participants were asked to read two paragraphs each from two Word documents using a specific screen magnification modality. Two types of documents were considered: a 1-column document, and a 3-column document. Text was displayed with a 9-point Helvetica font, with single line spacing and a whole blank line between paragraphs. The single-column document had 0.5" left and right margins. The three-column document had columns with width of 1.67" and spacing of 0.5" between columns. In the first two trials, participants accessed the single-column document, first with full screen magnification (first two paragraphs) then with screen lens magnification (next two paragraphs). They repeated the same sequence in the last two trials, this time on the three-column document.

3.2 Results

General Observations. All participants (except for the one who was excluded from the experiment, as mentioned earlier) were able to successfully complete all trials. The magnification factor α chosen by the participants ranged from 2.25 (P1) to 11.4 (P2, P3, P4). trials. The chosen lens size varied from 455x192 pixels (P1) to 1366x308 pixels (P3). The aspect ratio (width/height) of the lens varied from 1.2 (P2) to 4.4 (P3).

Reading Speed. When computing reading speeds, we considered the total number of standard-length words in the paragraphs being read, where the number of standard-length words in a paragraph is defined to be the total number of characters (including spaces and punctuation) divided by 6 [Carver 1990]. We observed a large variation in reading speed, from 14 words per minute (P5, screen lens, 1 column) to 208 words per minute (P6, full screen, 3 columns). Analysis of the data using paired t-test shows that the average reading speed for the 3-column document using screen lens was faster than for the 1-column document using full screen magnification (p=0.03) or screen lens (p=0.04). In addition, the average reading speed for the 3-column document using full screen magnification was found to be significantly larger than for the 1-column document using the screen lens modality. The mean reading speeds measured over all trials for these 5 participants was 42.3 words/minute, which is consistent with what found in [8] (mouse mode, low vision subjects, 44 words/minute).

Gaze Tracking Quality. Even though all participants successfully completed the calibration phase, analysis of the data collected shows that gaze tracking was successful (as defined by an effective reading rate of 20 Hz or more on average) only for four of the six participants: P1, P2 (except for the first trial,) P4, and P6. No useful gaze data could be obtained for P3 and for P5 (except for one trial, wherein the reading rate for P5 reached 5 readings per second.)

Mouse Motion / Eye Gaze Tracks.
Figure 1 (left column) shows the recorded mouse (FoM) tracks and gaze tracks, superimposed on the un-magnified screen, for a set of representative trials. The figure also shows the plots of the X- and Y-coordinate of the gaze point samples as a function of time, as well as of the location of the element looked at in the un-magnified screen. Not surprisingly, gaze points are located close to the FoM in the case of screen lens (SL) magnification. This is clearly seen in the plots, and confirmed by the moderate ($\rho \geq 0.4$) to strong ($\rho \geq 0.6$) correlation coefficients measured between gaze point and FoM, with the notable exception of P6 for the 3-column document. P6 chose to use a wide window with a relatively low magnification (2.59), such that the window contained the whole width of the magnified column, requiring almost no horizontal motion of the window

during reading. For what concerns correlation under full screen (FS) magnification, results varied from moderate correlation in the 1-column case for P1 and P6, to very weak correlation (|ρ|<0.2) in the other cases.

4 Study 2: Gaze-Based Control

4.1 Method

Participants. This study included two participants from Study 1 (P2, P6) and a new participant (P7), who was not part of Study 1 (see Table 1). Of note, P2 underwent the Study 2 experiment one year after the Study 1 experiment, while P6 did both experiments in the same day.

Apparatus. We developed two simple systems for gaze-based control of full screen magnification, and one system for screen lens control, as described below.

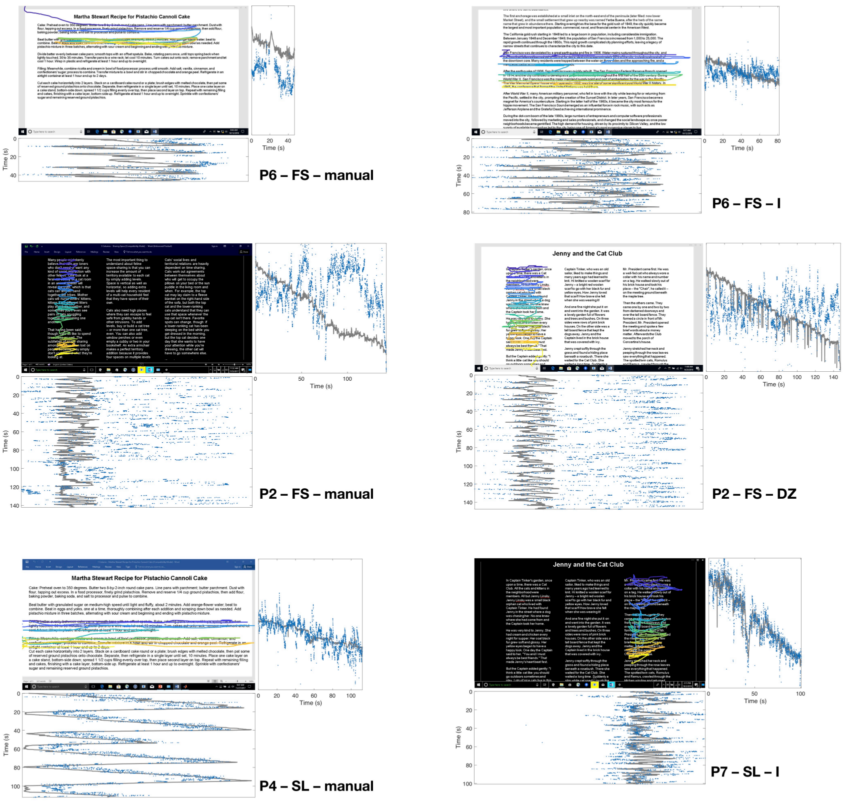

Figure 1. Sample data from our trials from different participants. The colored line on the (un-magnified) screenshots represents the track of the FoM (with color changing from purple to yellow as a function of time). The X- and Y-coordinate of the recorded gaze points are shown as blue dots at the bottom and to the right, respectively, of the screenshots. The same plots show the coordinates of $p(t)$, the location on the un-magnified screen of the element been looked at. Left column: manual control with full screen (FS) or screen lens (LS) magnification. Right column: gaze-contingent control using the FS-DZ, FS-I, or SL-I algorithms.

Full Screen–Dead Zone (FS–DZ). With this control modality, the onscreen content is scrolled only when the user's gaze point is outside of a central rectangular region (*dead zone*). Eight *scroll zones* are defined bordering the dead zone (the scroll zones are invisible to the user). When gaze fixates on a scroll zone, the onscreen content is scrolled with constant velocity towards the opposite size of the screen. For example, if one is reading a line of magnified text, and reaches the rightmost scroll zone, the FoM is moved to the right (remember from Eq. (2) that the screen content appears to move in the opposite direction of the FoM). The onscreen text thus moves to the left, making more magnified content available within the viewport for reading. In order to move to the beginning of the next line, one needs to look intently at the left edge of the screen, causing the FoM to move to the left and the magnified content to the right. In the implementation used for our tests, the dead zone was set to be small, with horizontal and vertical sizes equal to 1/10 of the corresponding screen size. In practice, this meant that the screen content was scrolled most of the time. The horizontal and/or vertical component of the FoM velocity (when gaze was in a scroll zone) was set equal to $600/\alpha$ pixels per second (where α is the magnification factor). The only exception was when gaze falls on the leftmost scroll zone, in which case the (horizontal) velocity was doubled. This was done to facilitate moving to the beginning of the new line, which requires full scroll of the screen content to the right.

Full Screen–Integrative (FS–I). Inspired by classic control theory, this mechanism implements an integrative controller. The general idea is to move the FoM such that the user, while reading text, is led to "naturally" gaze at a fixed location, that is chosen to be the center of the screen. If $g(t)$ is the location of the gaze point at time t, and $m(t)$ is the location of the FoM, the algorithm moves the FoM with velocity $v_m(t)$ defined by: $e(t) = g(t) - s;\ v_m(t) = \gamma e(t)$ where s is the location of the center of the screen, and γ is a positive coefficient set to $0.1/\alpha$. If gaze remains fixed at the center of the screen, the FoM also remains static. As soon as one moves their gaze to the right (e.g., while reading a line of text), the FoM also moves right, effectively scrolling the screen content to the left, with a speed that depends on the distance of the gaze point to the center of the screen. In order to reduce the risk of continuous motion due to small saccades (which could lead to motion sickness [9]), the error term is checked against a threshold (i.e., the X- or Y- component of $e(t)$ is set to 0 if its magnitude is smaller than a positive constant ϵ_X or ϵ_Y). In our experiments, we set ϵ_X and ϵ_Y equal to 1/20 of the width and height of the screen, respectively. This effectively creates a dead zone identical to that of the FS-DZ algorithm. The main difference between the two is that, while the velocity in a scroll zone is constant for FS-DZ, it can be controlled in FS-I by moving one's gaze closer or farther away from the screen center.

Screen Lens–Integrative (SL–I). For the case of screen lens magnification, we implemented an algorithm identical to FS–I, with the critical difference that the FoM $m(t)$ is made to smoothly move towards the current gaze point $g(t)$, rather than towards the center of the screen. This is obtained by simply replacing $e(t)$ as defined above with $g(t) - m(t)$. For this case, the extent of the dead zone is set to a much smaller value.

Experiment. The experiment was conducted in a very similar way to Study 1. The same computer and gaze tracker were used, with the difference that a mouse was not made available, as the participants were tasked with reading magnified text only using gaze-based control. Participants attempted to read two paragraphs from the 1-column document using FS-DZ, then the next two paragraph using FS-I, and the final two paragraph using SL-I. The process was then repeated for the 3-column document.

4.2 Results

All three participants were able to use both systems for gaze-based full screen magnification control (FS–DZ and FS–I) without any particular difficulty (although P2 struggled while reading the 3-column document under FS-I). However, gaze-based control for the screen lens modality (SL–I) was found to be very challenging. Only P6 was able to complete the trials on both documents, while P7 was only successful in the 3-column document. P2 was not able to complete either trial with SL-I. The reading speed was within the range of those recorded for mouse-based control, except for P6 using SL-I, in which case the reading speed was substantially lower than for the equivalent trials with mouse control. Failure to use SL-I appeared to be caused by over-compensation when the lens was not centered where desired, which often resulted in loss of control.

All three participants complained of some fatigue and of a somewhat "unnatural" reading experience while using gaze-based control. No participant felt motion sickness. Sample mouse and gaze tracks are shown in Figure 1 (right column).

5 Conclusions

We presented a preliminary study on the feasibility of a system that relies on the user's gaze direction to control the focus of magnification of a screen magnifier. This analysis may inform the future design of a gaze-contingent magnification control.

While our Study 2 showed that gaze-contingent magnification control is feasible (at least in the full screen modality), more research is needed on the design of effortless, ergonomic, and natural gaze-based control mechanisms. Based on our experience with this system, we believe that gaze-based magnification control should be built around two critical components: (1) a predictor of the element p in the un-magnified screen the user is interested in looking at; and (2) a mechanism to decide the location $\hat{p}$ where to map this element after magnification. From these two values, an appropriate location m for the FoM can be derived. For what concerns the second component (finding an appropriate location $\hat{p}$ for the magnified element), different strategies are available. For example, one may choose to maintain an almost stable gaze location, while moving the FoM $m(t)$ such that the desired text position $p(t)$ at all times falls, after magnification, on the same or similar screen location $\hat{p}$: $m(t) = (\alpha p(t) - \hat{p})/(\alpha - 1)$, where α is the magnification factor. In this case, one may expect little correlation between gaze point and FoM. At the other end of the spectrum, one may control the FoM such that one's gaze is led to follow the text line exactly as it would without magnification, i.e. $g(t) = p(t)$. This can be obtained by ensuring that the FoM always falls on the location in the un-magnified screen of the text element currently being gazed at ($m(t) = p(t) = g(t)$). From our analysis of the gaze tracks vis-à-vis the mouse tracks from our Study

1, no patterns emerged supporting either mechanism, suggesting that our participants chose control strategies that are in the middle ground between these two extremes. Ultimately, any control mechanism needs to be validated by proper user studies, which should include qualitative subjective measures besides standard quantitative metrics such as reading speed.

References

1. Aguilar, C., and E. Castet. 2012. Use of a gaze-contingent augmented-vision aid to improve reading with central field loss. Investigative Ophthalmology & Visual Science 53 (14): 4390.
2. Aguilar, C., and E. Castet. 2017. Evaluation of a gaze-controlled vision enhancement system for reading in visually impaired people. PLoS One 12 (4).
3. Ashmore, M., A. Duchowski, and G. Shoemaker. 2005. Efficient eye pointing with a fisheye lens. Proceedings of Graphics interface 2005.
4. Beckmann, P, and G Legge. 1996. Psychophysics of reading - XIV. The page navigation problem in using magnifiers. Vision Research 36 (22): 3723-3733.
5. Bruggeman, H., and G. Legge. 2002. Psychophysics of Reading-XIX. Hypertext Search and Retrieval With Low Vision. Proceedings of the IEEE 90 (1): 94-103.
6. Carver, R.P. 1990. Reading Rate: A Review of Research and Theory. San Diego: Academic Press.
7. Hallett, E., D. Wayne, T. Jewett, and K. Vu. 2017. How Screen Magnification with and Without Word-Wrapping Affects the User Experience of Adults with Low Vision. International Conference on Applied Human Factors and Ergonomics (AHFE 2017).
8. Harland, S., G. Legge, and A. Luebker. 1998. Psycophysics of Reading: XVII. Low-vision performances with four types of electronically magnified text. Optometry and Vision Science 75 (3): 183-190.
9. Hoeft, R., W. Buff, E. Cook, K. Stanney, and S. Wilson. 2002. Improving Assistive Technologies for the Visually Impaired: Minimizing the Side Effects of Magnification Products. Proceedings of the Human Factors and Ergonomics Society Annual Meeting.
10. Maus, N., D. Rutledge, S. Al-Khazraji, R. Bailey, C. Ovesdotter Alm, and K. Shinohara. 2020. Gaze-guided Magnification for Individuals with Vision Impairments. In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems.
11. Pölzer, S., A. Gander, and K. Miesenberger. 2018. Gaze based magnification to assist visually impaired persons. In International Conference on Computers Helping People with Special Needs., 2018.
12. Schwarz, T., Akbarioroumieh, A., Melfi, G. and Stiefelhagen, R. 2020. September. Developing a Magnification Prototype Based on Head and Eye-Tracking for Persons with Low Vision. In International Conference on Computers Helping People with Special Needs, 2020
13. Wallis, T., M. Dorr, and P. Bex. 2015. Sensitivity to gaze-contingent contrast increments in naturalistic movies: An exploratory report and model comparison. Journal of vision 15 (8): 3.
14. Werblin, F., R. Massof, N. Ross, D. Natale, C. Bradley, B. Yuval, and D. Teitelbaum. 2015. Gaze-Directed Magnification: Developing a Head-Mounted, Wide Field, Immersive Electronic Low Vision Aid. ARVO Annual Meeting Abstract.