TOUCHPAD MAPPER: Examining Information Consumption From 2D Digital Content Using Touchpads by Screen-Reader Users

Ather Sharif*
asharif@cs.washington.edu
Paul G. Allen School of Computer Science & Engineering | DUB Group, University of Washington
Seattle, Washington, USA

Venkatesh Potluri*
vpotluri@cs.washington.edu
Paul G. Allen School of Computer Science & Engineering | CREATE, University of Washington
Seattle, Washington, USA

Jazz R. X. Ang
jazz.angruixia@gmail.com
Human-Centered Design & Engineering | DUB Group, University of Washington
Seattle, Washington, USA

Jacob O. Wobbrock
wobbrock@uw.edu
The Information School | DUB Group, University of Washington
Seattle, Washington, USA

Jennifer Mankoff
Paul G. Allen School of Computer Science & Engineering | CREATE, University of Washington
Seattle, Washington, USA

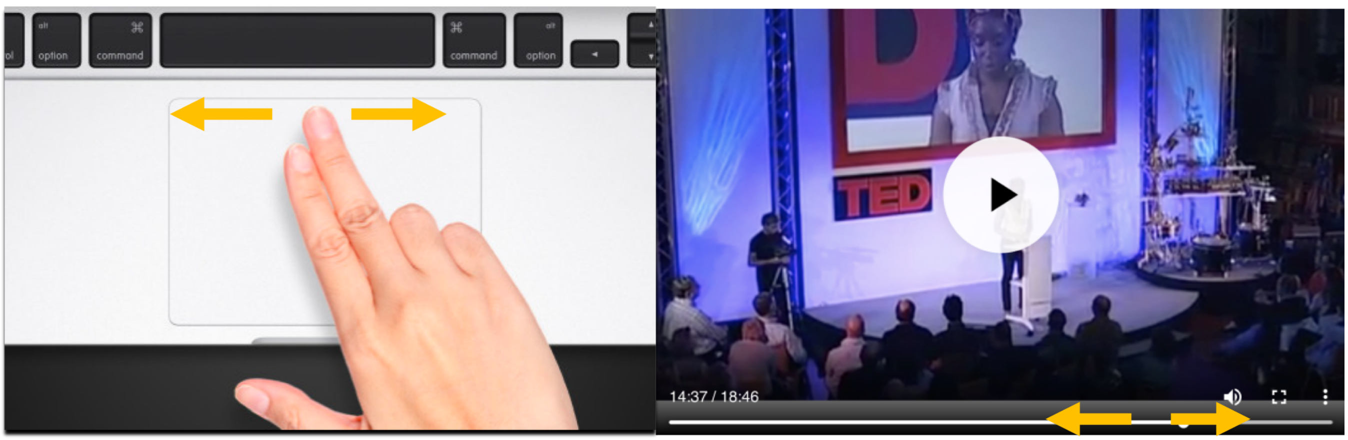

Figure 1: Using the touchpad to control a video seek bar (Usage Scenario 2).

Abstract

Touchpads are used widely to interact with computers, yet they provide minimal utility for screen-reader users. We explore the utility of touchpads as input devices for screen-reader users through the development and preliminary evaluation of TOUCHPAD MAPPER. This system maps digital content (i.e., images and videos) to the physical coordinates of a touchpad. We examined two usage scenarios: (1) identification of objects and their relative positioning in an image and (2) controlling a video seek bar and slider with rewinding and fast-forwarding features. We conducted task-based semi-structured interviews with two screen-reader users to assess our system's performance. The participants reported positive experiences, highlighting that they extracted information faster using our system than the conventional keyboard-only interaction.

CCS Concepts

• **Human-centered computing → Accessibility systems and tools**; **Visualization toolkits**; **User interface toolkits**.

Keywords

touchpad, screen reader, blind, interaction, non-visual, video

*These authors contributed equally to this work.

ASSETS '24, October 27–30, 2024, St. John's, NL, Canada

ACM ISBN 979-8-4007-0677-6/24/10
https://doi.org/10.1145/3663548.3688505

ACM Reference Format:
Ather Sharif, Venkatesh Potluri, Jazz R. X. Ang, Jacob O. Wobbrock, and Jennifer Mankoff. 2024. TOUCHPAD MAPPER: Examining Information Consumption From 2D Digital Content Using Touchpads by Screen-Reader Users. In *The 26th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '24), October 27–30, 2024, St. John's, NL, Canada.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3663548.3688505

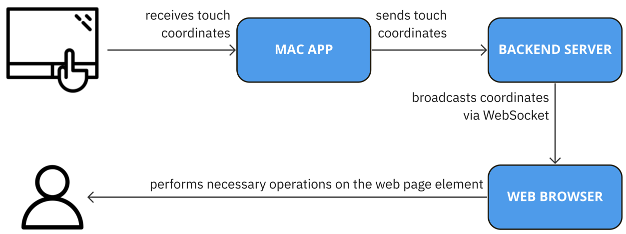

Figure 2: System diagram of Touchpad Mapper.

Figure 3: Image containing several landmarks (Scenario 1).

1 Introduction

Touchpads have existed as part of laptop computers for several decades and provide utility to users through multi-touch gesture support, input capabilities, and vibrotactile feedback. They are also commonly utilized as standalone devices. However, this widely used hardware offers minimal utility to screen-reader users for meaningfully interacting with digital content. Consequently, the digital interactions of screen-reader users remain predominantly linear and through discrete input devices, such as keyboards.

Exploratory efforts [2, 4, 7, 8, 11] have demonstrated the potential of non-linear information extraction using desktop screen readers. For example, Kane *et al.* [6] developed `Slide Rule`, which allows users to interact with mobile touch-screen-based devices by using gestural input and audio output. Similarly, JAWS [9], NVDA [1], and VoiceOver [5], among other screen readers, support exploring the interface by mousing (similar to smartphone screen reader interactions). While these efforts are plausible, they did not explore the usage of touchpads. In contrast, Touchpad Mapper demonstrates the *untapped* potential of touchpads for mapping 2D digital content, such as images and videos, to the physical layout of the touchpad. To our knowledge, Touchpad Mapper is the first demonstration of touchpad-enabled interaction techniques for screen-reader users.

We present Touchpad Mapper, a system that enables screen-reader users to utilize a touchpad to interact with our two usage scenarios: images and videos. Specifically, Touchpad Mapper maps an image to the physical dimensions of a touchpad using absolute coordinates, enabling the identification of objects via computer vision algorithms and their relative positioning in the image through finger-touch navigation. Additionally, Touchpad Mapper supports users to control a video seek bar and slider to rewind or fast-forward the video using the touchpad.

We assessed the usability of our system through preliminary task-based semi-structured interviews with two screen-reader users, examining the utility of our system's interactions. Participants found our system to be useful and faster than a conventional keyboard interaction with digital content. They shared positive experiences and suggestions for improvement.

2 Touchpad Mapper

We present Touchpad Mapper, its usage scenarios for images and videos, and its system design (see Figure 2).

2.0.1 Scenario 1: Images. In this scenario, we mapped an image to the dimensions of the touchpad, assisting users to identify objects in the image by moving their finger within the physical bounds of the touchpad, similar to the interaction `Slide Rule` enables on smartphones [6]. By mapping the absolute dimensions, our system accounts for the varying sizes of touchpads on different devices. We used Google's Vision API[1], , which uses advance artificial intelligence models in conjunction with object localization, to identify objects and their locations within the image on page load. When a user moves over a recognized object in an image, Touchpad Mapper reads the object's label aloud using the Web Speech API[2]. For example, in Figure 3, a user could locate landmarks, such as "Eiffel Tower" and "Statue of Liberty", and their position relative to each other. Without using our system, this interaction is not possible for screen-reader users; they rely on creators to provide this information via alt-text, which is often missing or insufficient [10, 12].

2.0.2 Scenario 2: Videos. For this scenario, we mapped the length of a video seek bar to the top horizontal edge of the touchpad (see Figure 1), enabling users to alter the seek position. The absolute mapping assists users in calibrating the seek position more effectively than through keyboard-based interactions. For example, if the user wishes to navigate to the middle of the video, given that the left edge represents the start and the right edge represents the end, the user can physically determine the seek position by touching the midway point between the start and end. Compared to this interaction, keyboard interactions would require linear navigation to locate the seek bar and continuous pressing of a designated key to advance the video by predefined increments.

We used the bottom horizontal edge of the touchpad for the video slider to rewind and fast-forward the video. Positioning the finger toward the left rewinds the video, whereas placing it toward the right fast-forwards the video. Additionally, we added four levels of rewinding and fast-forwarding, calculated using the distance from the mid-point: 5 seconds, 10 seconds, 30 seconds, and 60 seconds. For example, if the user places their finger in the first quartile between the mid-point and right edge, they can fast-forward with an increment of 5 seconds. Placing the finger in the second, third, and fourth quartile would have an increment of 10, 30, and 60 seconds, respectively.

[1] https://cloud.google.com/vision
[2] https://developer.mozilla.org/en-US/docs/Web/API/Web_Speech_API

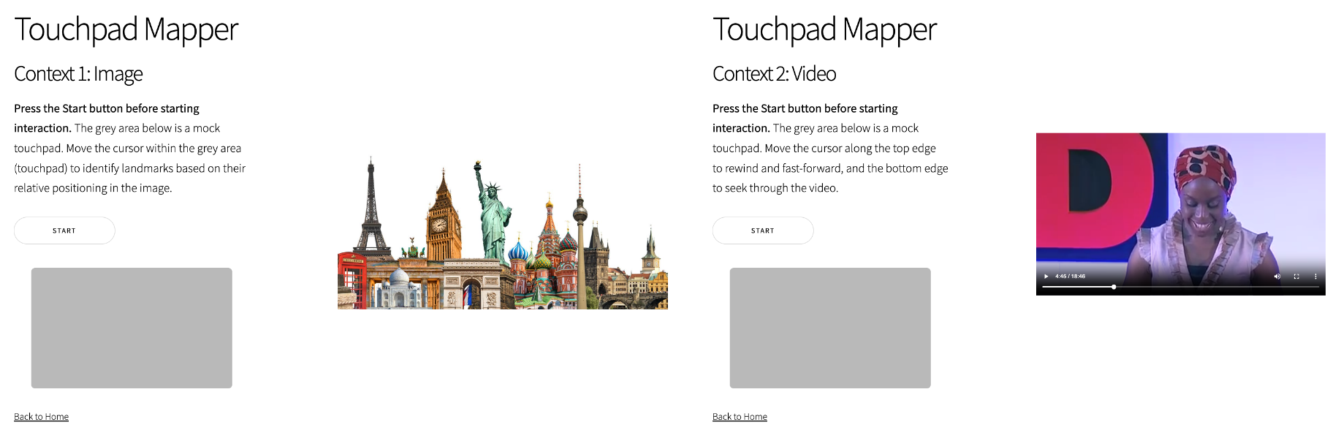

Figure 4: Screen capture of the mock demonstration of our two usage scenarios: image (left) and video (right). Image is Usage Scenario 1 and video is Usage Scenario 2.

2.1 Acquiring Touch Coordinates

To create Touchpad Mapper, we developed a MacOS App using Apple's AppKit. We utilized the NSEvent[3] API to acquire touch coordinates and gestures from touchpad events. To send this information to the browser client, we created a WebSocket.[4] We developed a backend server using NodeJS to receive the coordinates from our App and broadcast this data to the user's browser.

2.2 Browser Client

To explore the utility of non-visual touchpad interactions, we created two web pages, one with an image and one with a video. We used JavaScript and the WebSocket API to receive touch coordinates from the backend server. Each page contained a client-side script to determine and perform actions based on the coordinates and relay information to the user with the SpeechSynthesis interface of the Web Speech API. For example, on the image page, we used the coordinates to announce the objects in the image.

2.3 Mock Demonstration for Non-Screen-Reader Users

As Touchpad Mapper requires an installation on users' computer systems, we created a web application as a mock demonstration for non-screen-reader users. Specifically, we built a website using ReactJS[5] and hosted it using Github Pages. The website contains a dedicated web page for each usage scenario, displaying a mock touchpad and instructions to test our system. Specifically, our mock implementation allows users to move their mouse over the mock touchpad and hear the responses that screen-reader users would hear when interacting with their touchpads. Figure 4 shows a screen capture of the two web pages. Our mock demonstration can be found at https://athersharif.github.io/touchpadmapper/.

[3]An object that contains information about an input action, such as a keypress.
[4]A computer communications protocol that provides full-duplex communication channels over a single TCP connection.
[5]https://react.dev/

3 Task-Based Interviews

We conducted preliminary task-based semi-structured interviews with two blind screen-reader users to investigate the utility of touchpad interactions for screen-reader users. Our goal was to solicit their initial feedback and gather areas of improvement for future iterations of our system.

3.1 Participants

One participant identified as a completely blind woman, and the other participant was a man with pinhole vision. Their average age was 26 years. We recruited the participants through word-of-mouth and snowball sampling. None of the participants had prior familiarity with using touchpads. We compensated the participants with a $20 Amazon gift card for their time. As our goal was to conduct a preliminary assessment of Touchpad Mapper, we only recruited two participants.

3.2 Tasks & Analysis

After collecting the participants' consent and demographics, we first introduced our system and its features to the participants. Then, we asked them to perform two tasks using our system: (1) Explore Figure 3 freely to identify various landmarks (Usage Scenario 1) and (2) navigate a TED talk video to its midway point (Usage Scenario 2). Then, upon task completion, we conducted a semi-structured interview to solicit their feedback. We analyzed their responses using standard thematic analysis protocols [3].

3.3 Results

Our participants highlighted several beneficial use cases of utilizing touchpads ranging from mathematical content to web page elements, especially images and data visualizations. They reported positive experiences interacting with Touchpad Mapper and considered it faster than a conventional keyboard interaction. Participants also provided us with feedback on future improvements, such as enabling object detection to explore video frames—a feature our team is currently exploring.

4 Discussion & Conclusion

We investigated the utility of a touchpad as an input device for screen-reader users. We developed Touchpad Mapper, a system that maps the dimensions of an image to the absolute coordinates of a touchpad and enables image exploration via touch and object recognition. Touchpad Mapper also enables video exploration and playback control. We conducted task-based interviews with two screen-reader users, soliciting initial feedback and areas of improvement. They reported Touchpad Mapper to offer useful, efficient, and promising interaction techniques.

Given the preliminary nature of our work and our goal to solicit initial feedback on our system, we conducted task-based interviews with two screen-reader users. We plan on co-designing the next iteration of Touchpad Mapper with more participants, subsequently conducting usability and longitudinal studies for qualitative, quantitative, and subjective assessments. We also intend to use more usage scenarios in our future work, especially those widely present on the web, such as tables and date picker widgets. Additionally, our usage scenario with images was limited to landscape images; work is underway to support the mapping of portrait images. Through the development of Touchpad Mapper, we hope to uncover new avenues of information extraction other than the conventional keyboard-based interaction.

Acknowledgments

This work was supported by the University of Washington Center for Research and Education on Accessible Technology and Experiences (CREATE). We thank Matt Zeigler, Yisu Wang, James Fogarty, Jon Froehlich, and Richard Ladner for their comments and feedback. We remember our departed team member, Zoey. Without her feline support, the *purr*suit of this work would not have been as enjoyable. May she cross the rainbow bridge.

References

[1] NV Access. 2006. NV Access | Download. https://www.nvaccess.org/download/. (Accessed on 08/08/2021).

[2] Mark S. Baldwin, Gillian R. Hayes, Oliver L. Haimson, Jennifer Mankoff, and Scott E. Hudson. 2017. The Tangible Desktop: A Multimodal Approach to Nonvisual Computing. *ACM Transactions on Accessible Computing (TACCESS)* 10, 3 (Nov. 2017), 9. https://doi.org/10.1145/3075222

[3] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[4] Anhong Guo, Xiang'Anthony' Chen, Haoran Qi, Samuel White, Suman Ghosh, Chieko Asakawa, and Jeffrey P Bigham. 2016. Vizlens: A robust and interactive screen reader for interfaces in the real world. In *Proceedings of the 29th annual symposium on user interface software and technology*. 651–664.

[5] Apple Inc. 2005. Accessibility - Vision - Apple. https://www.apple.com/accessibility/vision/.

[6] Shaun K. Kane, Jeffrey P. Bigham, and Jacob O. Wobbrock. 2008. Slide Rule: Making Mobile Touch Screens Accessible to Blind People Using Multi-Touch Interaction Techniques. In *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility - Assets '08*. ACM Press, Halifax, Nova Scotia, Canada, 73. https://doi.org/10.1145/1414471.1414487

[7] Rushil Khurana, Duncan McIsaac, Elliot Lockerman, and Jennifer Mankoff. 2018. Nonvisual Interaction Techniques at the Keyboard Surface. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. ACM Press, Montreal QC, Canada, 1–12. https://doi.org/10.1145/3173574.3173585

[8] Barbara Leporini and Fabio Paternò. 2004. Increasing usability when interacting through screen readers. *Universal access in the information society* 3 (2004), 57–70.

[9] Freedom Scientific. 1995. JAWS® – Freedom Scientific. https://www.freedomscientific.com/products/software/jaws/. (Accessed on 08/08/2021).

[10] Ather Sharif, Sanjana S. Chintalapati, Jacob O. Wobbrock, and Katharina Reinecke. 2021. Understanding Screen-Reader Users' Experiences with Online Data Visualizations. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility* (Virtual Event, USA) *(ASSETS '21)*. Association for Computing Machinery, New York, NY, USA, Article 14, 16 pages. https://doi.org/10.1145/3441852.3471202

[11] Ather Sharif, Olivia H. Wang, Alida T. Muongchan, Katharina Reinecke, and Jacob O. Wobbrock. 2022. VoxLens: Making Online Data Visualizations Accessible with an Interactive JavaScript Plug-In. In *CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 478, 19 pages. https://doi.org/10.1145/3491102.3517431

[12] Jonathan Zong, Crystal Lee, Alan Lundgard, JiWoong Jang, Daniel Hajas, and Arvind Satyanarayan. 2022. Rich Screen Reader Experiences for Accessible Data Visualization. *Computer Graphics Forum* 41, 3 (2022), 15–27. https://doi.org/10.1111/cgf.14519 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.14519