

# *Leveraging Spatial Interactions to Enhance BLV Users' Navigation of Virtual Environments*

## *Project Proposal*

*By: Anna-Samsara Daefler, Demetria Mack (Group name: SpaceInt)*

*avd2131, dmm2285*

### ***Abstract***

The Computer-Enabled Abilities Laboratory, also known as CEAL, is a research team based at Columbia University led by Professor Brian Smith. Their work focuses on developing computers that help users perceive and interact with the world around them, perfectly addressing the need to build technology for the people, not as a hindrance, but to truly embody the essence of Human-Computer Interaction (HCI).

We want to highlight one particular project: their work on understanding and developing spatial interaction techniques for desktop readers. The team concentrated on providing assistance for blind and low-vision (BLV) users to navigate the web using desktop readers. Their approach supplemented desktop screen readers with layout-based (i.e., spatial) interactions—specifically, directional input and spatial sound output.

Our project aims to extend the technology they initiated in this project area ~~or in the event we cannot get access, we'd want to apply this to another domain.~~ The code for CEAL's spatial interactions project (WebNExt) is open source and available at <https://github.com/ColumbiaCEAL/spatial-interactions-extension>.

### ***Project Description***

~~(Notice we are still looking to define our project; here are two ideas we are interested in, hoping to get feedback on which would be better)~~

#### **~~Idea 1: Multimodal Feedback in Spatial Interaction~~**

~~Concept: Add real-time audio, visual, or haptic feedback to gesture interactions.~~

We would like to build in additional feedback to WebNExt around visual concepts such as page scrolling and when directional navigation hits the visual edge of the web page to make spatial browsing more seamless. Participants in the WebNExt study expressed interest in explicit confirmation when they reached the edge of a page especially in the scenario when the final element in a direction is not at the screen's edge but in the center of the page. Navigating past large blank regions where spatial browsing would make a large shift was also described as disconcerting for users without any explicit feedback. The additional feedback could be provided through special sound cues or descriptions read by the screen reader. We are planning to validate our system through unit testing.

As a stretch goal we are considering providing an overview of page layout using AI. In the WebNExt study participants mentioned wanting to quickly get a summarized understanding of the content on the screen i.e. where different types of content are placed.

### ***Additional Ideas***

**Note:** Here we would love feedback on if any of these ideas are better than the main idea we proposed above .

- Focus on one particular area such as collaborative work and how spatial interactions could add value. CollabAlly, for example, uses spatial audio to indicate a collaborator's location within a document. Many participants in the WebNExT study mentioned that the technology enhanced their collaboration with sighted users. We could focus on making it easier for BLV users to orient themselves in long collaborative-style documents.
- Add speech interactions, for example, to navigate to a certain location on the page or search for a word on the page.
- Use AI to describe images when alt text is missing.

**Tech Stack:** Python (MediaPipe for gesture detection), OpenCV for visual feedback, pydub or pygame for audio, optional Unity (C#) or Arduino for advanced visuals/haptics.

**Evaluation:** Compare user performance and satisfaction with vs. without feedback during task execution.

**Consideration/Drawbacks:** - requires us to interact with BLV participants, may not be able to gather enough initial feedback and insights in time for final presentation, getting permission from the lab participants to leverage their code

**How to build it (Still needs to be confirmed rough idea):**

- **Gesture Input:** Use the existing MediaPipe setup in Python to detect gestures.
- **Audio Feedback (Gamified):** Use pydub to play a sound when a gesture is triggered.
- **Visual Feedback:** Overlay simple animations using OpenCV or build a front end with PyQt/Tkinter for pop-up visuals.
- **Haptic Feedback:** Optional—connect to Arduino or game controller to trigger vibration.

**Idea 2:** Incorporate spatial interactions into another domain.

- We are considering gaming i.e. using spatial interactions to facilitate exploration for BLV players. Spatial sound output could enhance BLV player's immersion in games. Sighted players are driven to explore virtual environments through visual cues. This project would explore spatial audio cues as a way of encouraging discovery in a virtual environment.
- The lab also suggests that route finding could benefit from spatial interactions by allowing BLV users to evaluate multiple route options. Navigation information is usually presented linearly. I expect spatial interactions could help reduce cognitive load for BLV individuals navigating routes through richer spatial understanding and facilitate communication about routes with sighted people.

**Considerations:** Ideally, we would evaluate our system with the feedback of BLV participants. However, as mentioned in idea 1, this may be a challenge.

*Follow-up Steps*

- Reach out to the CEAL Lab team to request access to code files related to spatial interaction techniques for desktop readers project
  - Contact: [chheda@cs.washington.edu](mailto:chheda@cs.washington.edu)
- Determine: Can the full pipeline (gesture detection + multimodal feedback + evaluation) be completed within your semester timeline?
  - Determine: What specific usability metrics will we track?
    - \*It might be unnecessary as we are enhancing the technology based on the limitations set by the writers. The need for our work can be considered validated from the data they gathered during initial analyses with participants.
- Map out Feasibility timelines:
  - Will we be able to conduct studies with real users? With the possibility of leveraging some of the users from the CEAL labs that the team has worked with on previous projects?
  - Will conducting online usability testing be just as effective?

### **References:**

- <https://ceal.cs.columbia.edu/>
- <https://ceal.cs.columbia.edu/spatialinteractions/#project-publications>
- [https://en.wikipedia.org/wiki/Human%E2%80%93computer\\_interaction#:~:text=Human%E2%80%93computer%20interaction%20\(HCI\),with%20computers%20in%20novel%20ways.](https://en.wikipedia.org/wiki/Human%E2%80%93computer_interaction#:~:text=Human%E2%80%93computer%20interaction%20(HCI),with%20computers%20in%20novel%20ways.)
- <https://digitalcommons.lindenwood.edu/cgi/viewcontent.cgi?article=1721&context=faculty-research-papers>
- The ACM Reference Format for this paper is: Arnavi Chheda-Kothary, Ather Sharif, David A. Rios, and Brian A. Smith. 2025. “It Brought Me Joy”: Opportunities for Spatial Browsing in Desktop Screen Readers. In CHI Conference on Human Factors in Computing Systems (CHI ’25), April 26-May 1, 2025, Yokohama, Japan. ACM, New York, NY, USA, 18 pages. <https://doi.org/10.1145/3706598.3714125>.